B5 (8)

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 : C07K 15/10, A61K 39/36 — A1

(11) International Publication Number: WO 93/01213

(43) International Publication Date: 21 January 1993 (21.01.93)

(21) International Application Number: PCT/US92/05661

(22) International Filing Date: 10 July 1992 (10.07.92)

(30) Priority data:
729,134 — 12 July 1991 (12.07.91) — US
730,452 — 15 July 1991 (15.07.91) — US

(71) Applicant: IMMULOGIC PHARMACEUTICAL CORPORATION [US/US]; One Kendall Square, Building 600, Cambridge, MA 02139 (US).

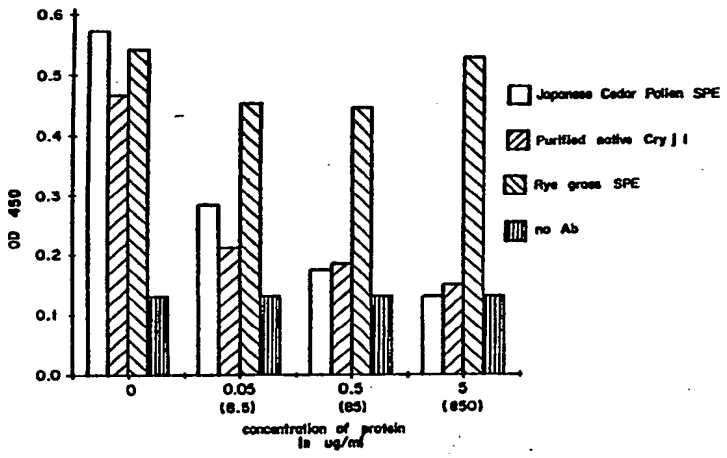
(72) Inventors: GRIFFITH, Irwin, J. ; 13 Southwick Road, North Reading, MA 01864 (US). POLLOCK, Joanne ; 51 Newcomb Street, Arlington, MA 02174 (US). BOND, Julian, F. ; 294 Commercial Street, Weymouth, MA 02188 (US).

(74) Agents: CHANNING, Stacey, L. et al.; Immulogic Pharmaceutical Corporation, Patent Department, One Kendall Square, Building 600, Cambridge, MA 02139 (US).

(81) Designated States: AU, CA, JP, KR, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LU, MC, NL, SE).

Published
With international search report.

(54) Title: ALLERGENIC PROTEINS AND PEPTIDES FROM JAPANESE CEDAR POLLEN



(57) Abstract

The present invention provides nucleic acid sequences coding for the *Cryptomeria japonica* major pollen allergen *Cry j* I and fragments thereof. The present invention also provides purified *Cry j* I and at least one fragment thereof produced in a host cell transformed with a nucleic acid sequence coding for *Cry j* I or at least one fragment thereof and fragments of *Cry j* prepared synthetically. *Cry j* I and fragments thereof are useful for diagnosing, treating, and preventing Japanese cedar pollinosis.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria
AU	Australia
BB	Barbados
BE	Belgium
BF	Burkina Faso
BG	Bulgaria
BJ	Benin
BR	Brazil
CA	Canada
CF	Central African Republic
CG	Congo
CH	Switzerland
CI	Côte d'Ivoire
CM	Cameroon
CS	Czechoslovakia
DE	Germany
DK	Denmark
ES	Spain
FI	Finland
FR	France
GA	Gabon
GB	United Kingdom
GN	Guinea
GR	Greece
HU	Hungary
IE	Ireland
IT	Italy
JP	Japan
KP	Democratic People's Republic of Korea
KR	Republic of Korea
LI	Liechtenstein
LK	Sri Lanka
LU	Luxembourg
MC	Monaco
MG	Madagascar
ML	Mali
MN	Mongolia
MR	Mauritania
MW	Malawi
NL	Netherlands
NO	Norway
PL	Poland
RO	Romania
RU	Russian Federation
SD	Sudan
SE	Sweden
SN	Senegal
SU	Soviet Union
TD	Chad
TG	Togo
US	United States of America

ALLERGENIC PROTEINS AND PEPTIDES FROM JAPANESE CEDAR POLLEN

Background of the Invention

Genetically predisposed individuals, who make up about 10% of the population, become hypersensitized (allergic) to antigens from a variety of environmental sources to which they are exposed. Those antigens that can induce immediate and/or delayed types of hypersensitivity are known as allergens. (King, T.P., *Adv. Immunol.* 23: 77-105, (1976)). Anaphylaxis or atopy, which includes the symptoms of hay fever, asthma, and hives, is one form of immediate allergy. It can be caused by a variety of atopic allergens, such as products of grasses, trees, weeds, animal dander, insects, food, drugs, and chemicals.

The antibodies involved in atopic allergy belong primarily to the IgE class of immunoglobulins. IgE binds to mast cells and basophils. Upon combination of a specific allergen with IgE bound to mast cells or basophils, the IgE may be cross- linked on the cell surface, resulting in the physiological effects of IgE-antigen interaction. These physiological effects include the release of, among other substances, histamine, serotonin, heparin, a chemotactic factor for eosinophilic leukocytes and/or the leukotrienes, C4, D4, and E4, which cause prolonged constriction of bronchial smooth muscle cells (Hood, L.E. et al. *Immunology* (2nd ed.), The Benjamin/Cumming Publishing Co., Inc. (1984)). These released substances are the mediators which result in allergic symptoms caused by a combination of IgE with a specific allergen. Through them, the effects of an allergen are manifested. Such effects may be systemic or local in nature, depending on the route by which the antigen entered the body and the pattern of deposition of IgE on mast cells or basophils. Local manifestations generally occur on epithelial surfaces at the location at which the allergen entered the body. Systemic effects can include anaphylaxis (anaphylactic shock), which is the result of an IgE-basophil response to circulating (intravascular) antigen.

Japanese cedar (Sugi; *Cryptomeria japonica*) pollinosis is one of the most important allergic diseases in Japan. The number of patients suffering from this disease is on the increase and in some areas, more than 10% of the population

are affected. Treatment of Japanese cedar pollinosis by administration of Japanese cedar pollen extract to effect hyposensitization to the allergen has been attempted. Hyposensitization using Japanese cedar pollen extract, however, has drawbacks in that it can elicit anaphylaxis if high doses are used, whereas when low doses are used to avoid anaphylaxis, treatment must be continued for several years to build up a tolerance for the extract.

The major allergen from Japanese cedar pollen has been purified and designated as Sugi basic protein (SBP) or *Cry j* I. This protein is reported to be a basic protein with a molecular weight of 41-50 kDa and a pI of 8.8. There appear to be multiple isoforms of the allergen, apparently due in part to differential glycosylation (Yasueda et al. (1983) *J. Allergy Clin. Immunol.* 71: 77-86; and Taniai et al. (1988) *FEBS Letters* 239: 329-332. The sequence of the first twenty amino acids at the N-terminal end of *Cry j* I and a sixteen amino acid internal sequence have been determined (Taniai *supra*).

A second allergen from Japanese cedar pollen having a molecular weight of about 37 kDa known as *Cry j* II has also been reported (Sakaguchi et al. (1990) *Allergy* 45: 309-312). This allergen was found to have no immunological cross- reactivity with *Cry j* I. Most patients with Japanese cedar pollinosis were found to have IgE antibodies to both *Cry j* I and *Cry j* II, however, sera from some patients reacted with only *Cry j* I or *Cry j* II.

In addition to hyposensitization of Japanese cedar pollinosis patients with low doses of Japanese cedar pollen extract, U.S. patent 4,939,239 issued July 3, 1990 to Matsuhashi et al. discloses a hyposensitization agent comprising a saccharide covalently linked to a Japanese cedar pollen allergen for hyposensitization of persons sensitive to Japanese cedar pollen. This hyposensitization agent is reported to enhance the production of IgG and IgM antibodies, but reduce production of IgE antibodies which are specific to the allergen and responsible for anaphylaxis and allergy. The allergens used in the hyposensitization agent preferably have an NH_2-terminal amino acid sequence of Asp-Asn-Pro-Ile-Asp-Ser-X-Trp-Arg-Gly-Asp-Ser-Asn-Trp-Ala-Gln-Asn-Arg-Met-Lys-, wherein X is Ser, Cys, Thr, or His (SEQ ID NO: 18). Additionally, Usui et al. (1990) *Int. Arch. Allergy Appl. Immunol.* 91: 74-79 reported that the ability of a Sugi basic protein (i.e., *Cry j* I)-pullulan conjugate to elicit the Arthus reaction was markedly reduced, about 1,000 times lower than that of native Sugi basic protein and

suggested that the Sugi basic protein-pullulan conjugate would be a good candidate for desensitization therapy against cedar pollinosis.

The *Cry j* I allergen found in *Cryptomeria japonica* has also been found to be cross-reactive with allergens in the pollen from other species of trees, including *Cupressus sempervirens*. Panzani et al. (*Annals of Allergy* 57: 26-30 (1986)) reported that cross reactivity was detected between allergens in the pollens of *Cupressus sempervirens* and *Cryptomeria japonica* in skin testing, RAST and RAST inhibition. A 50 kDa allergen isolated from Mountain Cedar (*Juniperus sabinoides*) has the NH_2-terminal sequence AspAsnProIleAsp (SEQ ID NO: 25) (Gross et al, (1978) Scand. J. Immunol. 8: 437-441) which is the same sequence as the first five amino acids of the NH-2 terminal end of the *Cry j* I allergen. The *Cry j I* allergen has also been found to be allergenically cross-reactive with the following species of trees: *Cupressus arizonica, Cupressus macrocarpa, Juniperus virginiana, Juniperus communis, Thuya orientalis*, and *Chamaecyparis obtusa*.

Despite the attention Japanese cedar pollinosis allergens have received, definition or characterization of the allergens responsible for its adverse effects on people is far from complete. Current desensitization therapy involves treatment with pollen extract with its attendant risks of anaphylaxis if high doses of pollen extract are administered, or long desensitization times when low doses of pollen extract are administered.

Summary of the Invention

The present invention provides nucleic acid sequences coding for the *Cryptomeria japonica* major pollen allergen *Cry j* I and fragments thereof. The present invention also provides purified *Cry j* I and at least one fragment thereof produced in a host cell transformed with a nucleic acid sequence coding for *Cry j* I or at least one fragment thereof and fragments of *Cry j I* prepared synthetically. As used herein, a fragment of the nucleic acid sequence coding for the entire amino acid sequence of *Cry j* I refers to a nucleotide sequence having fewer bases than the nucleotide sequence coding for the entire amino acid sequence of *Cry j* I and/or mature *Cry j* I. *Cry j* I and fragments thereof are useful for diagnosing, treating, and preventing Japanese cedar pollinosis. This invention is more particularly described in the appended claims and is described in its preferred embodiments in the following description.

Brief Description of the Drawings

Fig. 1a is a graphic representation of affinity purified Cry j I on Superdex 75 (2.6 by 60 cm) equilibrated with 10 mM sodium acetate (pH 5.0) and 0.15 M NaCl;

Fig. 1b shows an SDS-PAGE (12.5%) analysis of the fractions from the major peak shown in Fig 1a;

Fig. 2 shows a Western blot of isoforms of purified native Cry j I proteins separated by SDS-PAGE and probed with mAB CBF2;

Fig. 3 is a graphic representation of allergic sera titration of different purified fractions of purified native *Cry j* I using plasma from a pool of fifteen allergic patients;

Figs. 4a-b show the composite nucleic acid sequence from the two overlapping clones JC 71.6 and pUC19JC91A coding for *Cry j* I. The complete cDNA sequence for *Cry j* I is composed of 1312 nucleotides, including 66 nucleotides of 5' untranslated sequence, an open reading frame starting with the codon for an initiating methionine, of 1122 nucleotides, and a 3' untranslated region. Figs. 4a-b also show the deduced amino acid sequence of *Cry j* I;

Fig. 5a is a graphic representation of the results of IgE binding reactivity wherein the coating antigen is soluble pollen extract (SPE) from Japanese cedar pollen;

Fig. 5b is a graphic representation of the results of IgE binding reactivity wherein the coating antigen is purified native *Cry j* I;

Fig. 6 is a graphic representation of the results of a competition ELISA with pooled human plasma (PHP) from 15 patients wherein the coating antigen is soluble pollen extract (SPE) from Japanese cedar pollen;

Fig. 7 is a graphic representation of the results of a competition ELISA using plasma from individual patients (indicated by patient numbers) wherein the coating antigen is soluble pollen extract (SPE) from Japanese cedar pollen and the competing antigen is purified native *Cry j* I;

Fig. 8a is a graphic representation of the results from a direct binding ELISA using plasma from seven individual patients (indicated by patient numbers) wherein the coating antigen is soluble pollen extract (SPE) from Japanese cedar

pollen:

Fig. 8b is a graphic representation of the results from a direct binding ELISA using plasma from seven individual patients (indicated by patient numbers) wherein the coating antigen is denatured soluble pollen extract which has been denatured by boiling in the presence of a reducing agent, DTT;

Fig. 9 is a graphic representaion of a direct ELISA where the wells were coated with recombinant *Cry j* I (r*Cry j* I) and IgE binding was assayed on individual patients;

Fig. 10a is a graphic representation of the results of a capture ELISA using pooled human plasma from fifteen patients wherein the wells were coated with CBF2 (IgG) mAb, PBS was used as a negative antigen control, and the antigen was purified recombinant *Cry j* I;

Fig. 10b is a graphic representation of the results of a capture ELISA using rabbit anti-*Amb a*I and II, wherein the wells were coated with 20 µg/ml CBF2 (IgG), PBS was used as a negative antigen control and the antigen was purified recombinant *Cry j* I;

Fig. 11 is a graphic representation of a histamine release assay performed on one Japanese cedar pollen allergic patient using SPE from Japanese cedar pollen, purified native *Cry j* I and recombinant *Cry j* I as the added antigens; and

Fig. 12 is a graphic representation of the results of a T cell proliferation assay using blood from patient #999 wherein the antigen is recombinant *Cry j* I protein, purified native *Cry j* I protein, or recombinant *Amb a* 1.1.

Detailed Description of the Invention

The present invention provides nucleic acid sequences coding for *Cry j* I, the major allergen found in Japanese cedar pollen. The nucleic acid sequence coding for *Cry j* I preferably has the sequence shown in Figs. 4a and 4b (SEQ ID NO: 1). The nucleic acid sequence coding for *Cry j* I shown in Figs. 4a and 4b (SEQ ID NO: 1) contains a 21 amino acid leader sequence from base 66 through base 128. This leader sequence is cleaved from the mature protein which is encoded by bases 129 through 1187. The deduced amino acid sequence of *Cry j* I is also shown in Figs. 4a and 4b (SEQ ID NO: 2). The nucleic acid sequence of the invention codes for a protein having a predicted molecular weight of 38.5 kDa, with

a pI of 7.8, and five potential N-linked glycosylation sites. Utilization of these glycosylation sites will increase the molecular weight and affect the pI of the mature protein. The deduced amino acid sequence for the mature protein encoded by the nucleic acid sequence of the invention is identical with the known NH_2-terminal and internal amino acid sequences reported by Taniai et al., *supra*. The NH_2-terminal end of *Cry j* I reported by Taniai et al., *supra* has the sequence shown in SEQ ID NO: 18. The internal sequence reported by Taniai et al., *supra* has the sequence GluAlaPheAsnValGluAsnGlyAsnAlaThrProGlnLeuThrLys (SEQ ID NO: 19). There are sequence polymorphisms observed in the nucleic acid sequence of the invention. For example, single independent nuleotide substitutions at the codons encoding amino acids 38, 51 and 74 (GGA vs. GAA, GTG vs. GCG, and GGG vs. GAG, respectively) of SEQ ID #1 may result in amino acid polymorphisms (G vs. E, V vs. A, and G vs. E, respectively) at these sites. In addition, a single nucleotide substitution has been detected in one cDNA clone derived from *Cryptomeria japonica* pollen collected in Japan. This substitution in the codon for amino acid 60 (TAT vs. CAT) of SEQ ID #1 may result in an amino acid polymorphism (Y vs. H) at this site. Additional silent nucleotide substitutions have been detected. It is expected that there are additional sequence polymorphisms, and it will be appreciated by one skilled in the art that one or more nucleotides (up to about 1% of the nucleotides) in the nucleic acid sequence coding for *Cry j* I may vary among individual *Cryptomeria japonica* plants due to natural allelic variation. Any and all such nucleotide variations and resulting amino acid polymorphisms are within the scope of the invention. Furthermore, there may be one or more family members of *Cry j* I. Such family members are defined as proteins related in function and amino acid sequence to *Cry j* I but encoded by genes at separate genetic loci.

Fragments of the nucleic acid sequence coding for fragments of *Cry j* I are also within the scope of the invention. Fragments within the scope of the invention include those coding for parts of *Cry j* I which induce an immune response in mammals, preferably humans, such as stimulation of minimal amounts of IgE; binding of IgE; eliciting the production of IgG and IgM antibodies; or the eliciting of a T cell response such as proliferation and/or lymphokine secretion and/or the induction of T cell anergy. The foregoing fragments of *Cry j* I are referred to herein as antigenic fragments. Fragments within the scope of the invention also include those capable of hybridizing with nucleic acid from other plant species for use in

screening protocols to detect allergens that are cross-reactive with *Cry j* I. As used herein, a fragment of the nucleic acid sequence coding for *Cry j* I refers to a nucleotide sequence having fewer bases than the nucleotide sequence coding for the entire amino acid sequence of *Cry j* I and/or mature *Cry j* I. Generally, the nucleic acid sequence coding for the fragment or fragments of *Cry j* I will be selected from the bases coding for the mature protein, however, in some instances it may be desirable to select all or a part of a fragment or fragments from the leader sequence portion of the nucleic acid sequence of the invention. The nucleic acid sequence of the invention may also contain linker sequences, modified restriction endonuclease sites and other sequences useful for cloning, expression or purification of *Cry j* I or fragments thereof.

A nucleic acid sequence coding for *Cry j* I may be obtained from *Cryptomeria japonica* plants. However, Applicants have found that mRNA coding for *Cry j* I could not be obtained from commercially available *Cryptomeria japonica* pollen. This inability to obtain mRNA from the pollen may be due to problems with storage or transportation of commercially available pollen. Applicants have found that fresh pollen and staminate cones are a good source of *Cry j* I mRNA. It may also be possible to obtain the nucleic acid sequence coding for *Cry j* I from genomic DNA. *Cryptomeria japonica* is a well-known species of cedar, and plant material may be obtained from wild, cultivated, or ornamental plants. The nucleic acid sequence coding for *Cry j* I may be obtained using the method disclosed herein or any other suitable techniques for isolation and cloning of genes. The nucleic acid sequence of the invention may be DNA or RNA.

The present invention provides expression vectors and host cells transformed to express the nucleic acid sequences of the invention. Nucleic acid coding for *Cry j* I, or at least one fragment thereof may be expressed in bacterial cells such as *E. coli*, insect cells (baculovirus), yeast, or mammalian cells such as Chinese hamster ovary cells (CHO). Suitable expression vectors, promoters, enhancers, and other expression control elements may be found in Sambrook et al. *Molecular Cloning: A Laboratory Manual*, second edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (1989). Other suitable expression vectors, promoters, enhancers, and other expression elements are known to those skilled in the art. Expression in mammalian, yeast or insect cells leads to partial or complete glycosylation of the recombinant material and formation of any inter- or intra-chain

disulfide bonds. Suitable vectors for expression in yeast include YepSec1 (Baldari et al. (1987) *Embo J.* 6: 229-234); pMFa (Kurjan and Herskowitz (1982) *Cell* 30: 933-943); JRY88 (Schultz et al. (1987) *Gene* 54: 113-123) and pYES2 (Invitrogen Corporation, San Diego, CA). These vectors are freely available. Baculovirus and mammalian expression systems are also available. For example, a baculovirus system is commercially available (PharMingen, San Diego, CA) for expression in insect cells while the pMSG vector is commerically available (Pharmacia, Piscataway, NJ) for expression in mammalian cells.

For expression in *E. coli*, suitable expression vectors include, among others, pTRC (Amann et al. (1988) *Gene* 69: 301-315); pGEX (Amrad Corp., Melbourne, Australia); pMAL (N.E. Biolabs, Beverly, MA); pRIT5 (Pharmacia, Piscataway, NJ); pET-11d (Novagen, Madison, WI) Jameel et al., (1990) *J. Virol.* 64:3963-3966; and pSEM (Knapp et al. (1990) *BioTechniques* 8: 280-281). The use of pTRC, and pET-11d, for example, will lead to the expression of unfused protein. The use of pMAL, pRIT5 pSEM and pGEX will lead to the expression of allergen fused to maltose E binding protein (pMAL), protein A (pRIT5), truncated ß-galactosidase (PSEM), or glutathione S-transferase (pGEX). When *Cry j* I, fragment, or fragments thereof is expressed as a fusion protein, it is particularly advantageous to introduce an enzymatic cleavage site at the fusion junction between the carrier protein and *Cry j* I or fragment thereof. *Cry j* I or fragment thereof may then be recovered from the fusion protein through enzymatic cleavage at the enzymatic site and biochemical purification using conventional techniques for purification of proteins and peptides. Suitable enzymatic cleavage sites include those for blood clotting Factor Xa or thrombin for which the appropriate enzymes and protocols for cleavage are commercially available from for example Sigma Chemical Company, St. Louis, MO and N.E. Biolabs, Beverly, MA. The different vectors also have different promoter regions allowing constitutive or inducible expression with, for example, IPTG induction (PRTC, Amann et al., (1988) *supra*: pET-11d, Novagen, Madison, WI) or temperature induction (pRIT5, Pharmacia, Piscataway, NJ) . It may also be appropriate to express recombinant *Cry j* I in different *E. coli* hosts that have an altered capacity to degrade recombinantly expressed proteins (e.g. U.S. patent 4,758,512). Alternatively, it may be advantageous to alter the nucleic acid sequence to use codons preferentially utilized by *E. coli*, where such nucleic acid alteration would not affect the amino acid

sequence of the expressed protein.

Host cells can be transformed to express the nucleic acid sequences of the invention using conventional techniques such as calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection, or electroporation. Suitable methods for transforming the host cells may be found in Sambrook et al. *supra*, and other laboratory textbooks.

The nucleic acid sequences of the invention may also be synthesized using standard techniques.

The present invention also provides a method of producing purified Japanese cedar pollen allergen *Cry j* I or at least one fragment thereof comprising the steps of culturing a host cell transformed with a DNA sequence encoding Japanese cedar pollen allergen *Cry j* I or at least one fragment thereof in an appropriate medium to produce a mixture of cells and medium containing said Japanese cedar pollen allergen *Cry j* I or at least one fragment thereof; and purifying the mixture to produce substantially pure Japanese cedar pollen allergen *Cry j* I or at least one fragment thereof. Host cells transformed with an expression vector containing DNA coding for *Cry j* I or at least one fragment thereof are cultured in a suitable medium for the host cell. *Cry j* I protein and peptides can be purified from cell culture medium, host cells, or both using techniques known in the art for purifying peptides and proteins including ion-exchange chromatography, gel filtration chromatography, ultrafiltration, electrophoresis and immunopurification with antibodies specific for *Cry j* I or fragments thereof. The terms isolated and purified are used interchangeably herein and refer to peptides, protein, protein fragments, and nucleic acid sequences substantially free of cellular material or culture medium when produced by recombinant DNA techniques, or chemical precursors or other chemicals when synthesized chemically.

Another aspect of the invention provides preparations comprising Japanese cedar pollen allergen *Cry j* I or at least one fragment thereof synthesized in a host cell transformed with a DNA sequence encoding all or a portion of Japanese cedar pollen allergen *Cry j* I, or chemically synthesized, and purified Japanese cedar pollen allergen *Cry j* I protein, or at least one antigenic fragment thereof produced in a host cell transformed with a nucleic acid sequence of the invention, or chemically synthesized. In preferred embodiments of the invention the *Cry j* I protein is produced in a host cell transformed with the nucleic acid sequence coding for at least

the mature *Cry j* I protein.

Fragments of an allergen from Japanese cedar pollen, preferably *Cry j* I, eliciting a desired antigenic response (referred to herein as antigenic fragments) may be obtained, for example, by screening peptides recombinantly produced from the corresponding fragment of the nucleic acid sequence of the invention coding for such peptides, synthesized chemically using techniques known in the art, or produced by chemical cleavage of the allergen, the allergen may be arbitrarily divided into fragments of a desired length with no overlap of the peptides, or preferably divided into fragments of a desired length with no overlap of the peptides, or preferably divided into overlapping fragments of a desired length. The fragments are tested to determine their antigenicity (e.g. the ability of the fragment to induce an immune response). If fragments of Japanese cedar pollen allergen, e.g.*Cry j*, I are to be used for therapeutic purposes, then the fragments of Japanese cedar pollen allergen which are capable of eliciting a T cell response such as stimulation (i.e., proliferation or lymphokine secretion) and/or are capable of inducing T cell anergy are particularly desirable and fragments of Japanese cedar pollen which have minimal IgE stimulating activity are also desirable. Additionally, for therapeutic purposes, purifed Japanese cedar pollen allergens, e.g. *Cry j* I, and fragments thereof preferably do not bind IgE specific for Japanese cedar pollen or bind such IgE to a substantially lesser extent than the purified native Japanese cedar pollen allergen binds such IgE. If the purified Japanese cedar pollen allergen or fragment or fragments thereof bind IgE, it is preferable that such binding does not result in the release of mediators (e.g. histamines) from mast cells or basophils. Minimal IgE stimulating activity refers to IgE stimulating activity that is less than the amount of IgE production stimulated by the native *Cry j* I protein.

Purified protein allergens from Japanese cedar pollen or preferred antigenic fragments thereof, when administered to a Japanese cedar pollen-sensitive individual, or an individual allergic to an allergen cross-reactive with Japanese cedar pollen allergen, such as allergen from the pollen of *Cupressus sempervirens* or *Juniperus sabinoides* etc. (discussed previously) are capable of modifying the allergic response of the individual to Japanese cedar pollen or such cross-reactive allergen of the individual, and preferably are capable of modifying the B-cell response, T-cell response or both the B-cell and the T-cell response of the individual to the allergen. As used herein, modification of the allergic response of an

individual sensitive to a Japanese cedar pollen allergen can be defined as non-responsiveness or diminution in symptoms to the allergen, as determined by standard clinical procedures (See e.g. Varney et al, *British Medical Journal*, 302:265-269 (1990)).

The purified *Cry j* I protein or fragments thereof are preferably tested in mammalian models of Japanese cedar pollinosis such as the mouse model disclosed in Tamura et al. (1986) *Microbiol. Immunol.* 30: 883-896, or U.S. patent 4,939,239; or the primate model disclosed in Chiba et al. (1990) *Int. Arch. Allergy Immunol.* 93: 83-88. Initial screening for IgE binding to the protein or fragments thereof may be performed by scratch tests or intradermal skin tests on laboratory animals or human volunteers, or in *in vitro* systems such as RAST (radioallergosorbent test), RAST inhibition, ELISA assay, radioimmunoassay (RIA), or histamine release (see Examples 7 and 8).

Antigenic fragments of the present invention which have T cell stimulating activity, and thus comprise at least one T cell epitope are particularly desirable. T cell epitopes are believed to be involved in initiation and perpetuation of the immune response to a protein allergen which is responsible for the clinical symptoms of allergy. These T cell epitopes are thought to trigger early events at the level of the T helper cell by binding to an appropriate HLA molecule on the surface of an antigen presenting cell and stimulating the relevant T cell subpopulation. These events lead to T cell proliferation, lymphokine secretion, local inflammatory reactions, recruitment of additional immune cells to the site, and activation of the B cell cascade leading to production of antibodies. One isotype of these antibodies, IgE, is fundamentally important to the development of allergic symptoms and its production is influenced early in the cascade of events, at the level of the T helper cell, by the nature of the lymphokines secreted. A T cell epitope is the basic element or smallest unit of recognition by a T cell receptor, where the epitope comprises amino acids essential to receptor recognition. Amino acid sequences which mimic those of the T cell epitopes and which modify the allergic response to protein allergens are within the scope of this invention.

Exposure of patients to purified protein allergens of the present invention or to the antigenic fragments of the present invention which comprise at least one T cell epitope and are derived from protein allergens may tolerize or anergize appropriate T cell subpopulations such that they become unresponsive to

the protein allergen and do not participate in stimulating an immune response upon such exposure. In addition, administration of the protein allergen of the invention or an antigenic fragment of the present invention which comprises at least one T cell epitope may modify the lymphokine secretion profile as compared with exposure to the naturally-occurring protein allergen or portion thereof (e.g. result in a decrease of IL-4 and/or an increase in IL-2). Furthermore, exposure to such antigenic fragment or protein allergen may influence T cell subpopulations which normally participate in the response to the allergen such that these T cells are drawn away from the site(s) of normal exposure to the allergen (e.g., nasal mucosa, skin, and lung) towards the site(s) of therapeutic administration of the fragment or protein allergen. This redistribution of T cell subpopulations may ameliorate or reduce the ability of an individual's immune system to stimulate the usual immune response at the site of normal exposure to the allergen, resulting in a dimunution in allergic symptoms.

The purified *Cry j* I protein, and fragments or portions derived therefrom (peptides) can be used in methods of diagnosing, treating and preventing allergic reactions to Japanese cedar pollen allergen or a cross reactive protein allergen. Thus the present invention provides therapeutic compositions comprising purified Japanese cedar pollen allergen *Cry j* I or at least one fragment thereof produced in a host cell transformed to express *Cry j* I or at least one fragment thereof, and a pharmaceutically acceptable carrier or diluent. The therapeutic compositions of the invention may also comprise synthetically prepared *Cry j* I or at least one fragment thereof and a pharmaceutically acceptable carrier or diluent. Administration of the therapeutic compositions of the present invention to an individual to be desensitized can be carried out using known techniques. *Cry j* I protein or at least one fragment thereof may be administered to an individual in combination with, for example, an appropriate diluent, a carrier and/or an adjuvant. Pharmaceutically acceptable diluents include saline and aqueous buffer solutions. Pharmaceutically acceptable carriers include polyethylene glycol (Wie et al. (1981) *Int. Arch. Allergy Appl. Immunol.* 64:84-99) and liposomes (Strejan et al. (1984) *J. Neuroimmunol* 7: 27). For purposes of inducing T cell anergy, the therapeutic composition is preferably administered in nonimmunogenic form, e.g. it does not contain adjuvant. Such compositions will generally be administered by injection (subcutaneous, intravenous, etc.), oral administration, inhalation, transdermal

application or rectal administration. The therapeutic compositions of the invention are administered to Japanese cedar pollen-sensitive individuals at dosages and for lengths of time effective to reduce sensitivity (i.e. reduce the allergic response) of the individual to Japanese cedar pollen. Effective amounts of the therapeutic compositions will vary according to factors such as the degree of sensitivity of the individual to Japanese cedar pollen, the age, sex, and weight of the individual, and the ability of the *Cry j* I protein or fragment thereof to elicit an antigenic response in the individual.

The *Cry j* I cDNA (or the mRNA from which it was transcribed) or a portion thereof can be used to identify similar sequences in any variety or type of plant and thus, to identify or "pull out" sequences which have sufficient homology to hybridize to the *Cry j* I cDNA or mRNA or portion thereof, for example, DNA from allergens of *Cupressus sempervirens, Juniperus sabinoides etc.*, under conditions of low stringency. Those sequences which have sufficient homology (generally greater than 40%) can be selected for further assessment using the method described herein. Alternatively, high stringency conditions can be used. In this manner, DNA of the present invention can be used to identify, in other types of plants, preferably related families, genera, or species such as *Juniperus*, or *Cupressus*, sequences encoding polypeptides having amino acid sequences similar to that of Japanese cedar pollen allergen *Cry j* I, and thus to identify allergens in other species. Thus, the present invention includes not only *Cry j* I, but also other allergens encoded by DNA which hybridizes to DNA of the present invention. The invention further includes isolated allergenic proteins or fragments thereof that are immunologically related to *Cry j* I or fragments thereof, such as by antibody cross-reactivity wherein the isolated allergenic proteins or fragments thereof are capable of binding to antibodies specific for the protein and peptides of the invention, or by T cell cross-reactivity wherein the isolated allergenic proteins or fragments thereof are capable of stimulating T cells specific for the protein and peptides of this invention.

Proteins or peptides encoded by the cDNA of the present invention can be used, for example as "purified" allergens. Such purified allergens are useful in the standardization of allergen extracts which are key reagents for the diagnosis and treatment of Japanese cedar pollinosis. Furthermore, by using peptides based on the nucleic acid sequences of *Cry j* I, anti-peptide antisera or monoclonal antibodies can be made using standard methods. These sera or monoclonal antibodies can be

used to standardize allergen extracts.

Through use of the peptides and protein of the present invention, preparations of consistent, well-defined composition and biological activity can be made and administered for therapeutic purposes (e.g. to modify the allergic response of a Japanese cedar sensitive individual to pollen of such trees). Administration of such peptides or protein may, for example, modify B-cell response to *Cry j* I allergen, T-cell response to *Cry j* I allergen or both responses. Purified peptides can also be used to study the mechanism of immunotherapy of *Cryptomeria japonica* allergy and to design modified derivatives or analogues useful in immunotherapy.

Work by others has shown that high doses of allergens generally produce the best results (i.e., best symptom relief). However, many people are unable to tolerate large doses of allergens because of allergic reactions to the allergens. Modification of naturally-occurring allergens can be designed in such a manner that modified peptides or modified allergens which have the same or enhanced therapeutic properties as the corresponding naturally-occurring allergen but have reduced side effects (especially anaphylactic reactions) can be produced. These can be, for example, a protein or peptide of the present invention (e.g., one having all or a portion of the amino acid sequence of *Cry j* I), or a modified protein or peptide, or protein or peptide analogue. It is possible to modify the structure of a protein or peptide of the invention for such purposes as increasing solubility, enhancing therapeutic or preventive efficacy, or stability (e.g., shelf life ex vivo, and resistance to proteolytic degradation in vivo). A modified protein or peptide can be produced in which the amino acid sequence has been altered, such as by amino acid substitution, deletion, or addition, to modify immunogenicity and/or reduce allergenicity, or to which a component has been added for the same purpose. For example, the amino acid residues essential to T cell epitope function can be determined using known techniques (e.g., substitution of each residue and determination of the presence or absence of T cell reactivity). Those residues shown to be essential can be modified (e.g., replaced by another amino acid whose presence is shown to enhance T cell reactivity), as can those which are not required for T cell reactivity (e.g., by being replaced by another amino acid whose incorporation enhances T cell reactivity but does not diminish binding to relevant MHC). Another example of a modification of protein or peptides is substitution of cysteine residues preferably with alanine, serine, threonine, leucine or glutamic acid to minimize

dimerization via disulfide linkages. Another example of modification of the peptides of the invention is by chemical modification of amino acid side chains or cyclization of the peptide.

In order to enhance stability and/or reactivity, the protein or peptides of the invention can also be modified to incorporate one or more polymorphisms in the amino acid sequence of the protein allergen resulting from natural allelic variation. Additionally, D-amino acids, non-natural amino acids or non-amino acid analogues can be substituted or added to produce a modified protein or peptide within the scope of this invention. Furthermore, proteins or peptides of the present invention can be modified using the polyethylene glycol (PEG) method of A. Sehon and co-workers (Wie et al. *supra*) to produce a protein or peptide conjugated with PEG. In addition, PEG can be added during chemical synthesis of a protein or peptide of the invention. Modifications of proteins or peptides or portions thereof can also include reduction/ alyklation (Tarr in: *Methods of Protein Microcharacterization,* J.E. Silver ed. Humana Press, Clifton, NJ, pp 155-194 (1986)); acylation (Tarr, *supra*); chemical coupling to an appropriate carrier (Mishell and Shiigi, eds, *Selected Methods in Cellular Immunology,* WH Freeman, San Francisco, CA (1980); U.S. Patent 4,939,239; or mild formalin treatment (Marsh *International Archives of Allergy and Applied Immunology,* 41:199-215 (1971)).

To facilitate purification and potentially increase solubility of proteins or peptides of the invention, it is possible to add reporter group(s) to the peptide backbone. For example, poly-histidine can be added to a peptide to purify the peptide on immobilized metal ion affinity chromatography (Hochuli, E. et al., *Bio/Technology,* 6:1321-1325 (1988)). In addition, specific endoprotease cleavage sites can be introduced, if desired, between a reporter group and amino acid sequences of a peptide to facilitate isolation of peptides free of irrelevant sequences. In order to successfully desensitize an individual to a protein antigen, it may be necessary to increase the solubility of a protein or peptide by adding functional groups to the peptide or by not including hydrophobic T cell epitopes or regions containing hydrophobic epitopes in the peptides or hydrophobic regions of the protein or peptide.

To potentially aid proper antigen processing of T cell epitopes within a peptide, canonical protease sensitive sites can be recombinantly or synthetically engineered between regions, each comprising at least one T cell epitope. For

example, charged amino acid pairs, such as KK or RR, can be introduced between regions within a peptide during recombinant construction of the peptide. The resulting peptide can be rendered sensitive to cathepsin and/or other trypsin-like enzymes cleavage to generate portions of the peptide containing one or more T cell epitopes. In addition, such charged amino acid residues can result in an increase in solubility of a peptide.

Site-directed mutagenesis of DNA encoding a peptide or protein of the invention (e.g. *Cry j* I or a fragment thereof) can be used to modify the structure of the peptide or protein by methods known in the art. Such methods may, among others, include PCR with degenerate oligonucleotides (Ho et al., *Gene*, 77:51-59 (1989)) or total synthesis of mutated genes (Hostomsky, Z. et al., *Biochem. Biophys. Res. Comm.*,161:1056-1063 (1989)). To enhance bacterial expression, the aforementioned methods can be used in conjunction with other procedures to change the eucaryotic codons in DNA constructs encoding protein or peptides of the invention to ones preferentially used in *E. coli*, yeast, mammalian cells, or other eukaryotic cells.

Using the structural information now available, it is possible to design *Cry j* I peptides which, when administered to a Japanese cedar pollen sensitive individual in sufficient quantities, will modify the individual's allergic response to Japanese cedar pollen. This can be done, for example, by examining the structure of *Cry j* I, producing peptides (via an expression system, synthetically or otherwise) to be examined for their ability to influence B-cell and/or T-cell responses in Japanese cedar pollen sensitive individuals and selecting appropriate peptides which contain epitopes recognized by the cells. In referring to an epitope, the epitope will be the basic element or smallest unit of recognition by a receptor, particularly immunoglobulins, histocompatibility antigens and T cell receptors where the epitope comprises amino acids essential to receptor recognition. Amino acid sequences which mimic those of the epitopes and which are capable of down regulating allergic response to *Cry j* I can also be used.

It is now also possible to design an agent or a drug capable of blocking or inhibiting the ability of Japanese cedar pollen allergen to induce an allergic reaction in Japanese cedar pollen sensitive individuals. Such agents could be designed, for example, in such a manner that they would bind to relevant anti-*Cry j* I IgEs, thus preventing IgE-allergen binding and subsequent mast cell degranulation.

Alternatively, such agents could bind to cellular components of the immune system, resulting in suppression or desensitization of the allergic response to *Cryptomeria japonica* pollen allergens. A non-restrictive example of this is the use of appropriate B- and T-cell epitope peptides, or modifications thereof, based on the cDNA/protein structures of the present invention to suppress the allergic response to Japanese cedar pollen. This can be carried out by defining the structures of B- and T-cell epitope peptides which affect B- and T-cell function in *in vitro* studies with blood components from Japanese cedar pollen sensitive individuals.

Protein, peptides or antibodies of the present invention can also be used for detecting and diagnosing Japanese cedar pollinosis. For example, this could be done by combining blood or blood products obtained from an individual to be assessed for sensitivity to Japanese cedar pollen with an isolated antigenic peptide or peptides of *Cry j* I, or isolated *Cry j* I protein, under conditions appropriate for binding of components in the blood (e.g., antibodies, T-cells, B- cells) with the peptide(s) or protein and determining the extent to which such binding occurs.

The present invention also provides a method of producing *Cry j* I or fragment thereof comprising culturing a host cell containing an expression vector which contains DNA encoding all or at least one fragment of *Cry j* I under conditions appropriate for expression of *Cry j* I or at least one fragment. The expressed product is then recovered, using known techniques. Alternatively, *Cry j* I or fragment thereof can be synthesized using known mechanical or chemical techniques.

The DNA used in any embodiment of this invention can be cDNA obtained as described herein, or alternatively, can be any oligodeoxynucleotide sequence having all or a portion of a sequence represented herein, or their functional equivalents. Such oligodeoxynucleotide sequences can be produced chemically or enzymatically, using known techniques. A functional equivalent of an oligonucleotide sequence is one which is 1) a sequence capable of hybridizing to a complementary oligonucleotide to which the sequence (or corresponding sequence portions) of SEQ ID NO: 1 or fragments thereof hybridizes, or 2) the sequence (or corresponding sequence portion) complementary to SEQ ID NO: 1, and/or 3) a sequence which encodes a product (e.g., a polypeptide or peptide) having the same functional characteristics of the product encoded by the sequence (or corresponding sequence portion) of SEQ ID NO: 1. Whether a functional equivalent must meet

one or both criteria will depend on its use (e.g., if it is to be used only as an oligoprobe, it need meet only the first or second criteria and if it is to be used to produce a *Cry j* I allergen, it need only meet the third criterion).

The invention is further illustrated by the following non-limiting examples.

Example 1

Purification of Native Japanese Cedar Pollen Allergen (*Cry j* I)

The following is a description of the work done to biochemically purify the major allergen, *Cry j* I in the native form. The purification was modified from published procedures (Yasueda et al., *J. Allergy Clin. Immunol.* 71:77, 1983).

100g of Japanese cedar pollen obtained from Japan (Hollister-Stier, Spokane, WA) was defatted in 1 L diethyl ether three times, the pollen was collected after filtration and the ether was dried off in a vacuum.

The defatted pollen was extracted at 4°C overnight in 2 L extraction buffer containing 50 mM tris-HCl, pH 7.8, 0.2 M NaCl and protease inhibitors in final concentrations: soybean trypsin inhibitor (2 µg/ml), leupeptin (1 µg/ml), pepstatin A (1 µg/ml) and phenyl methyl sulfonyl fluoride (0.17 mg/ml). The insoluble material was reextracted with 1.2 L extraction buffer at 4°C overnight and both extracts were combined together and depigmented by batch absorption with Whatman DE-52 DEAE cellulose (200 g dry weight) equilibrated with the extraction buffer.

The depigmented material was then fractionated by ammonium sulfate precipitation at 80% saturation (4°C), which removed much of the lower molecular weight material. The resultant partially purified *Cry j* I was either dialyzed in PBS buffer and used in T cell studies (see Example 6) or subjected to further purification as described below.

The enriched *Cry j* I material was then dialyzed against 50 mM Na-acetate, pH 5.0 at 4°C with 50 mM Na-acetate, pH 5.0 with protease inhibitors. The unbound material (basic proteins) was then applied to a 50 ml cation exchange column (Whatman CM-52) which was equilibrated at 4°C with 10 mM Na-acetate, pH 5.0 with protease inhibitors. *Cry j* I was eluted in the early fractions of a linear gradient 0.3 M NaCl. The enriched *Cry j* I material was lyophilized and was then purified by FPLC over a 300 ml Superdex 75 column (Pharmacia) at a flow rate of 30 ml/h in 10 mM Na-acetate, pH

5.0 at 25°C.

The purified *Cry j* I was further applied to FPLC S-Sepharose 16/10 column chromatography (Pharmacia) with a linear gradient of 0 - 1 M NaCl at 25°C. *Cry j* I eluted as the major peak was subjected to a second gel filtration chromatography. FPLC Superdex 75 column (2.6 by 60 cm)(Pharmacia, Piscataway, NJ) was eluted with a downward flow of 10 mM Na-acetate, pH 5.0 with 0.15 M NaCl at a flow rate of 30 ml/h at 25°C. Fig. 1a shows the chromatography on gel filtration. Only *Cry j* I was detected (Fig. 1b, lane 2 to lane 8). *Cry j* I was fractionated into 3 bands as analyzed by SDS-PAGE using silver staining (Fig. 1b) As shown in Fig. 1b, SDS PAGE (12.5%) analysis of the fractions from the major peak shown in Fig. 1a was performed under reducing conditions. The gel was silver stained using the silver staining kit from Bio-Rad. The samples in each lane were as follows: Lane 1, prestained standard proteins (Gibco BRL) including ovalbumin (43,000 kD), carbonic anhydrase (29,000 kD), and α-lactoglobulin (18,400 kD); lane 2, fraction 36 ; lane 3 fraction 37; lane 4 fraction 38; lane 5 fraction 39 ; lane 6 fraction 41, lane 7 fraction 43; and lane 8 fraction 44. All fractions are shown in Fig. 1a.

These proteins were also analyzed by Western blotting using mouse monoclonal antibody CBF2 (Fig. 2). As shown in Fig. 2, an aliquot of fraction 36 (lane 1), fraction 39, (lane 2) and fraction 43 (lane 3) purified from the Superdex 75 as shown in Fig. 1 was separated by SDS-PAGE, electroblotted onto nitrocelluslose and probed with mAB CBF2. Biotinlylated goat anti-mouse Ig was used for the second antibody and bound antibody was revealed by ^{125}I-streptavidin. The monoclonal CBF2 was raised against ragweed allergen *Amb a* I by Dr. D. Klapper (Chapel Hill, N. Carolina). Because of the homology between the *Amb a* I and *Cry j* I sequences, a number of antibodies raised against *Amb a* I were tested for reactivity with *Cry j* I. The results showed that CBF2 recognized denatured *Cry j* I as detected by ELISA and Western blotting. In addition, Western blotting also demonstrated that no other bands were detected by CBF2, other than *Cry j* I in the expected molecular weight range (Fig. 2). These results were consistent with the findings from protein sequencing. When fraction 44 and fraction 39 (Fig 1b) were subjected to N-terminal sequencing, only *Cry j* I sequence was detected.

In summary, three *Cry j* I isoforms of different molecular weight were purified from pollen extract. The molecular weights estimated by SDS-PAGE ranged from 40-35 kD under both reducing and non-reducing conditions. The isoelectric point

of these isoforms is approximately 9.5-8.6, with an average pI of 9.0. The N-terminal 20 amino acid sequence was the same in these 3 bands and was identical to previously published *Cry j* I sequence (Taniai et al. *supra*). The 3 isoforms are all recognized by monoclonal antibody CBF2 as shown in the allergic sera titration of different purified subfractions of Cry j I using a pool of fifteen allergic patient plasma. They all bind allergic patient IgE (Fig. 3). The difference in molecular weight and isoelectric point might in part be due to post-translational modification, e.g. glycosylation, phosphorylation or lipid content might be different in these isoforms. The possibility that these different isoforms might be due to protease degradation cannot be ruled out at present even though it is unlikely due to the fact that four different protease inhibitors were used during extraction and purification. The other possibility could be due to polymorphism in the gene or alternate splicing in the mRNA though only one major form of *Cry j* I protein has been detected in cDNA cloning studies (see Example 4).

Another approach which may be used to purify native *Cry j* I or recombinant *Cry j* I is immunoaffinity chromatography. This technique provides a very selective protein purification due to the specificity of the interaction between monoclonal antibodies and antigen. For the purpose of producing *Cry j* I-reactive monoclonal antibodies, female Balbl/c mice were obtained from Jackson Labs. Each mouse was initially immunized intraperitoneally with 70-100 µg purified native *Cry j* I, (>99% purity lower band, as shown in Fig. 1b), emulsified in Freund's complete adjuvant. One further intravenous injection of 10 µg purified native *Cry j* I in PBS was given 54 days after the initial injection. The spleen was removed 3 days later and myeloma fusion was conducted as described (Current Protocols in Immunology, 1991, Coligan et al, eds.) using the myeloma line SP2.0. The cells were cultured in 10% fetal calf serum (Hybrimax), hypoxanthine and azaserine and wells containing colonies of hybridoma cells were screened for antibody production using antigen-binding ELISA.

Cells from positive wells were cloned at three-tenths cell/well in 10% fetal calf serum (Hybrimax), hypoxanthine and positive clones were subcloned one more time in hypoxanthine medium . Capture ELISA (see Example 7) was used for secondary and tertiary screening. This assay offers the advantage that a clone that recognizes the native protein may be selected and thus may be useful for immunoaffinity purification. Thus, the mAbs will provide a useful tool in purification of *Cry j* I from pollen extracts. Similarly, monoclonal antibodies that bind to recombinant *Cry j* I can also be used for

immunoaffinity chromatography. In addition, the monoclonal antibodies generated may be useful for diagnostic purposes. It may also be possible to raise different mAbs that show some specificity towards these different isoforms of *Cry j* I and thus would provide a useful tool to characterize these isoforms.

Example 2

Attempted Extraction of RNA From Japanese Cedar Pollen

Multiple attempts were made to obtain RNA from commercially-available, non-defatted, *Cryptomeria japonica* (Japanese cedar) pollen (Hollister Stier, Seattle, WA). Initially, the method of Sambrook et al., *Molecular Cloning. A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (1989) was used in which the sample was suspended and lysed in 4 M guanidine buffer, ground under liquid nitrogen, and pelleted through 5.7 M cesium chloride by ultracentrifugation. Various amounts (3, 5 and 10 g) of pollen in varying amounts of guanidine lysis buffer (10 and 25 ml) were tried. Centrifugation through cesium resulted in viscous material in the bottom of the tube, from which it was not possible to recover an RNA pellet. Although it was possible to obtain RNA from defatted *Ambrosia artemisiifolia* (ragweed) pollen (Greer Laboratories, Lenior, NC) using this protocol, defatting the *Cryptomeria japonica* pollen with acetone before guanidine extraction also did not yield any RNA, as determined by absorbance at A_{260}.

An acid phenol extraction of RNA according to the method in *Sambrook* et al., *supra* was attempted from *Cryptomeria japonica* pollen. The pollen was ground and sheared in 4.5 M guanidine solution, acidified by addition of 2 M sodium acetate, and extracted with water-saturated phenol plus chloroform. After precipitation, the pellet was washed with 4 M lithium chloride, redissolved in 10 mM Tris/5 mM EDTA/1% SDS, chloroform extracted, and re-precipitated with NaCl and absolute ethanol. It was possible to extract *Ambrosia artemisiifolia* but not *Cryptomeria japonica* RNA with this procedure.

Next, 4 g of *Cryptomeria japonica* pollen was suspended in 10 ml extraction buffer (50 mM Tris, pH 9.0, 0.2 M NaCl, 10 mM Mg acetate and diethylpyrocarbonate (DEPC) to 0.1%), ground in a mortar and pestle on dry ice, transferred to a centrifuge tube with 1% SDS, 10 mM EDTA and 0.5% N- lauroyl

sarcosine, and the mixture was extracted five times with warm phenol. The aqueous phase was recovered after the final centrifugation, 2.5 vol. absolute ethanol was added, and the mixture was incubated overnight at 4°C. The pellet was recovered by centrifugation, resuspended in 1 ml dH_20 by heating to 65°C, and reprecipitated by the addition of 0.1 vol. 3 M Na acetate and 2.0 vol. of ethanol. No detectable RNA was recovered in the pellet as judged by absorbance at A_{260} and gel electrophoresis.

Finally, 500 mg of *Cryptomeria japonica* pollen was ground by mortar and pestle on dry ice and suspended in 5 ml of 50 mM Tris pH 9.0 with 0.2 M NaCl, 1 mM EDTA, 1% SDS that had been treated overnight with 0.1% DEPC, as previously described in Frankis and Mascarhenas (1980) *Ann. Bot.* 45: 595- 599. After five extractions with phenol/chloroform/isoamyl alcohol (mixed at 25:24:1), material was precipitated from the aqueous phase with 0.1 volume 3 M sodium acetate and 2 volumes ethanol. The pellet was recovered by centrifugation, resuspended in dH_20 and heated to 65°C to solubilize the precipitated material. Further precipitations with lithium chloride were not done. There was no detectable RNA recovered, as determined by absorbance at A_{260} and gel electrophoresis.

In summary, it has not been possible to recover RNA from the commercial pollen. It is not known whether the RNA has been degraded during storage or shipment, or whether the protocols used in this example did not allow recovery of extant RNA. However, RNA was recovered from fresh *Cryptomeria japonica* pollen and staminate cone samples. (See Example 3.)

Example 3

Extraction of RNA From Japanese Cedar Pollen and Staminate Cones and Cloning of *Cry j I*

Fresh pollen and staminate cone samples, collected from a single *Cryptomeria japonica* (Japanese cedar) tree at the Arnold Arboretum (Boston, MA), were frozen immediately on dry ice. RNA was prepared from 500 mg of each sample, essentially as described by Frankis and Mascarenhas, *supra*. The samples were ground by mortar and pestle on dry ice and suspended in 5 ml of 50 mM Tris pH 9.0 with 0.2 M NaCl, 1 mM EDTA, 1% SDS that had been treated overnight with 0.1% DEPC. After five extractions with phenol/chloroform/isoamyl alcohol (mixed at 25:24:1), the RNA was precipitated from the aqueous phase with 0.1

volume 2 M sodium acetate and 2 volumes ethanol. The pellets were recovered by centrifugation, resuspended in dH_20 and heated to 65°C for 5 min. Two ml of 4 M lithium chloride were added to the RNA preparations and they were incubated overnight at 0°C. The RNA pellets were recovered by centrifugation, resuspended in 1 ml dH_20, and again precipitated with 3 M sodium acetate and ethanol overnight. The final pellets were resuspended in 100 µl dH_20 and stored at -80°C.

First strand cDNA was synthesized from 8 µg flowerhead and 4 µg pollen RNA using a commercially available kit (cDNA synthesis systems kit, BRL, Gaithersburg, MD) with oligo dT priming according to the method of Gubler and Hoffman (1983) *Gene* 25: 263-269. An attempt was made to amplify cDNA encoding *Cry j* I using the degenerate oligonucleotide CP-1 (which has the sequence 5'-GATAATCCGATAGATAG-3', wherein T at position 3 can also be C; T at position 6 can also be C; G at position 9 can also be A,T, or C; A at position 12 can also be T, or C; T at position 15 can also be C; A at position 16 can also be T; and G at position 17 can also be C; SEQ ID NO: 3) and primers EDT and ED. Primer EDT has the sequence 5'-GGAATTCTCTAGACTGCAGGTTTTTTTTTTTTTT-3'(SEQ ID NO: 24). Primer ED has the sequence 5'-GGAATTCTCTAGACTGCAGGT-3' (SEQ ID NO: 23). CP-1 is the degenerate oligonucleotide sequence encoding the first six amino acids of the amino terminus (AspAsnProIleAspSer, amino acids 1-6 of SEQ ID NO: 1) of *Cry j* I. EDT will hybridize with the poly A tail of the gene. All oligonucleotides were synthesized by Research Genetics, Inc. Huntsville, AL. Polymerase chain reactions (PCR) were carried out using a commercially available kit (GeneAmp DNA Amplification kit, Perkin Elmer Cetus, Norwalk, CT) whereby 10 µl 10x buffer containing dNTPs was mixed with 1 µg of CP- 1 and 1 µg of ED/EDT primers (ED:EDT in a 3:1 M ratio), cDNA (3-5 µl of a 20 µl first strand cDNA reaction mix), 0.5 µl Amplitaq DNA polymerase, and distilled water to 100 µl.

The samples were amplified with a programmable thermal controller (MJ Research, Inc., Cambridge, MA). The first 5 rounds of amplification consisted of denaturation at 94°C for 1 minute, annealing of primers to the template at 45°C for 1.5 minutes, and chain elongation at 70°C for 2 minutes. The final 20 rounds of amplification consisted of denaturation as above, annealing at 55°C for 1.5 minutes, and elongation as above. Five percent (5 µl) of this initial amplification was then used in a secondary amplification with 1 µg each of CP-2 (which has the sequence

5'- GGGAATTCAATTGGGCGCAGAATGG-3' wherein T at position 11 can also be C; G at position 17 can also be A, T, or C; G at position 20 can also be A; T at position 23 can also be C; and G at position 24 can also be C) (SEQ ID NO: 4), a nested primer, and ED, as above. The sequence 5'-GGGAATTC-3' (bases 1 through 8 of SEQ ID NO: 4) in primer CP-2 represents an *Eco* R1 site added for cloning purposes; the remaining degenerate oligonucleotide sequence encodes amino acids 13-18 of *Cry j* I (AsnTrpAlaGlnAsnArg, amino acids 13 through 18 of SEQ ID NO: 1). Multiple DNA bands were resolved on a 1% GTG agarose gel (FMC, Rockport, ME), none of which hybridized with ^{32}P end- labeled probe CP-3 (SEQ ID NO: 5) in a Southern blot performed according to the method in Sambrook et al. *supra*. Therefore, it was not possible to select a specific *Cry j* I DNA band and this approach was not pursued. CP-3 has the sequence 5'-CTGCAGCCATTTTCIACATTAAA-3' wherein A at position 9 can also be G; T at position 12 can also be C; A at position 18 can also be G; and A at position 21 can also be G) (SEQ ID NO: 5). Inosine (I) is used at position 15 in place of G or A or T or C to reduce degeneracy (Knoth et al. (1988) *Nucleic Acids Res.* 16: 10932). The sequence 5'-CTGCAG-3' (bases 1 through 6 of SEQ ID NO: 5) in primer CP-3 represent a Pst I site added for cloning purposes; the remaining degenerate oligonucleotide sequence is the non-coding strand sequence corresponding to coding strand sequence encoding amino acids PheAsnValGluAsnGly (amino acids 327 through 332 of SEQ ID NO: 1) from the internal sequence of *Cry j* I.

A primary PCR was also performed on first-strand cDNA using CP-1 (SEQ ID NO: 3) and CP-3 (SEQ ID NO: 5), as above. A secondary PCR was performed using 5% of the primary reaction using CP-2 (SEQ ID NO: 4) and CP-3 (SEQ ID NO: 5). Again, multiple bands were observed, none of which could be specifically identified in a Southern blot as *Cry j* I, and this approach was also not pursued.

Double-stranded cDNA was then synthesized from approximately 4 µg (pollen) or 8 µg (flowerhead) RNA using a commercially available kit (cDNA Synthesis System kit, BRL, Gaithersburg, MD). After a phenol extraction and ethanol precipitation, the cDNA was blunted with T4 DNA polymerase (Promega, Madison, WI), and ligated to ethanol precipitated, self-annealed, AT (SEQ ID NO: 20) and AL (SEQ ID NO: 22) oligonucleotides for use in a modified Anchored PCR reaction, according to the method in Rafnar et al. (1991) *J. Biol. Chem.* 266: 1229-

1236; Frohman et al. (1990) *Proc. Natl. Acad. Sci. USA* 85: 8998-9002; and Roux et al. (1990) *BioTech.* 8: 48-57. Oligonucleotide AT has the sequence 5'-GGGTCTAGAGGTACCGTCCGATCGATCATT-3'(SEQ ID NO: 20) (Rafnar et al. *supra*). Oligonucleotide AL has the sequence 5'-AATGATCGATGCT-3' (SEQ ID NO: 22) (Rafnar et al. *Supra.* The amino terminus of *Cry j* I was amplified from the linkered cDNA (3 ul from a 20 µl reaction) with 1 µg each of oligonucleotides AP (SEQ ID NO: 21) and degenerate *Cry j* I primer CP-7 (which has the sequence 5'-TTCATICGATTCTGGGCCCA-3' wherein G at position 8 can also be T; A at position 9 can also be G; C at position 12 can also be T; and G at position 15 can also be A, T, or C)(SEQ ID NO: 6). Inosine (I) is used at position 6 in place of G or A or T or C to reduce degeneracy (Knoth et al. *supra*). The degenerate oligonucleotide CP-7 (SEQ ID NO: 6) is the non-coding strand sequence corresponding to coding strand sequence encoding amino acids 14-20 (TrpAlaGlnAsnArgMetLys) from the amino terminus of *Cry j* I (amino acids 14-20 of SEQ ID NO: 1). Oligonucleotide AP has the sequence 5'-GGGTCTAGAGGTACCGTCCG-3' (SEQ ID NO: 21).

The primary PCR reaction was carried out as described herein. Five percent (5 µl) of this initial amplification was then used in a secondary amplification with 1 µg each of AP (SEQ ID NO: 21) and degenerate *Cry j* I primer CP-8 (SEQ ID NO: 7) an internally nested *Cry j* I oligonucleotide primer, as described herein. Primer CP-8 has the sequence 5'-CCTGCAGCGATTCTGGGCCCAAATT-3' wherein G at position 9 can also be T; A at position 10 can also be G; C at position 13 can also be T; G at position 16 can also be A, T, or C; and A at position 23 can also be G)(SEQ ID NO: 7). The nucleotides 5'-CCTGCAG-3' (bases 1 through 7 of SEQ ID NO: 7) represent a *Pst* I restriction site added for cloning purposes. The remaining degenerate oligonucleotide sequence is the non-coding strand sequence corresponding to coding strand sequence encoding amino acids 13-18 of *Cry j* I (AsnTrpAlaGlnAsnArg, amino acids 13-18 of SEQ ID NO: 1) from the amino terminus of *Cry j* I. The dominant amplified product was a DNA band of approximately 193 base pairs, as visualized on an ethidium bromide (EtBr)-stained 3% GTG agarose gel.

Amplified DNA was recovered by sequential chloroform, phenol, and chloroform extractions, followed by precipitation at -20°C with 0.5 volumes of 7.5 ammonium acetate and 1.5 volumes of isopropanol. After precipitation and washing

with 70% ethanol, the DNA was simultaneously digested with *Xba* I and *Pst* I in a 15 µl reaction and electrophoresed through a preparative 3% GTG NuSieve low melt gel (FMC, Rockport, ME). The appropriate sized DNA band was visualized by EtBr staining, excised, and ligated into appropriately digested M13mp18 for sequencing by the dideoxy chain termination method (Sanger et al. (1977) *Proc. Natl Acad Sci. USA* 74: 5463-5476) using a commercially available sequencing kit (Sequenase kit, U.S. Biochemicals, Cleveland, OH). It was initially thought that ligatable material could only be derived from staminate cone-derived RNA. However, upon subsequent examination, it was shown that ligatable material could be recovered from PCR product generated from pollen-derived RNA, and from staminate cone-derived RNA.

The clone designated JC71.6 was found to contain a partial sequence of *Cry j* I. This was confirmed as an authentic clone of *Cry j* I by having complete identity to the disclosed NH_2-terminal sequence of *Cry j* I (Taniai et al. *supra*). The amino acid at position 7 was determined to be cysteine (Cys) in agreement with the sequence disclosed in U.S. patent 4, 939,239. Amino acid numbering is based on the sequence of the mature protein; amino acid 1 corresponds to the aspartic acid (Asp) disclosed as the NH_2-terminus of *Cry j* I (Taniai et al. *supra*) The initiating methionine was found to be amino acid -21 relative to the first amino acid of the mature protein. The position of the initiating methionine was supported by the presence of upstream in-frame-stop codons and by 78% homology of the surrounding nucleotide sequence with the plant consensus sequence that encompasses the initiating methionine, as reported by Lutcke et al. (1987) *EMBO J.* 6:43-48.

The cDNA encoding the remainder of *Cry j* I gene was cloned from the linkered cDNA by using oligonucleotides CP-9 (which has the sequence 5'-ATGGATTCCCCTTGCTTA-3')(SEQ ID NO: 8) and AP (SEQ ID NO: 21) in the primary PCR reaction. Oligonucleotide CP-9 (SEQ ID NO: 8) encodes amino acids MetAspSerProCysLeu of *Cry j* I (amino acids -21 through -16 of SEQ ID NO: 1) from the leader sequence of *Cry j* I, and is based on the nucleotide sequence determined for the partial *Cry j* I clone JC76.1.

A secondary PCR reaction was performed on 5% of the initial amplification mixture, with 1 µg each of AP (SEQ ID NO: 21) and CP-10 (which has the sequence 5'-GGGAATTCGATAATCCCATAGACAGC-3')(SEQ ID NO: 9).

the nested primer. The nucleotide sequence 5'-GGGAATTC-3' of primer CP- 10 (bases 1 through 8 of SEQ ID NO: 9) represent an *Eco* RI restriction site added for cloning purposes. The remaining oligonucleotide sequence encodes amino acids 1-6 of *Cry j* I (AspAsnProIleAspSer) (amino acids 1 through 6 of SEQ ID NO: 1), and is based on the nucleotide sequence determined for the partial *Cry j* I clone JC76.1. The amplified DNA product was purified and precipitated as above, followed by digestion with *Eco* RI and *Xba* I and electrophoresis through a preparative 1% low melt gel. The dominant DNA band was excised and ligated into M13mp19 and pUC19 for sequencing. Again, ligatable material was recovered from cDNA generated from pollen-derived RNA, and from staminate cone-derived RNA. Two clones, designated pUC19JC91a and pUC19JC91d, were selected for full-length sequencing. They were subsequently found to have identical sequences.

DNA was sequenced by the dideoxy chain termination method (Sanger et al. *supra*) using a commercially available kit (sequenase kit (U.S. Biochemicals, Cleveland, OH). Both strands were completely sequenced using M13 forward and reverse primers (N.E. Biolabs, Beverly, MA) and internal sequencing primers CP-13 (SEQ ID NO: 10), CP-14 (SEQ ID NO: 11), CP-15 (SEQ ID NO: 12), CP-16 (SEQ ID NO: 13), CP-18 (SEQ ID NO: 15), CP-19 (SEQ ID NO: 16), and CP-20 (SEQ ID NO: 17). CP-13 has the sequence 5'-ATGCCTATGTACATTGC-3' (SEQ ID NO: 10). CP-13 (SEQ ID NO: 10) encodes amino acids 82-87 of *Cry j* I (MetProMetTyrIleAla, amino acids 82 through 87 of SEQ ID NO: 1). CP-14 has the sequence 5'-GCAATGTACATAGGCAT-3' (SEQ ID NO: 11) and corresponds to the non-coding strand sequence of CP-13 SEQ ID NO: 10). CP-15 has the sequence 5'- TCCAATTCTTCTGATGGT-3' ((SEQ ID NO: 12) which encodes amino acids 169-174 of *Cry j* I (SerAsnSerSerAspGly, amino acids 169 through 174 of SEQ ID NO: 1). CP-16 has the sequence 5'-TTTTGTCAATTGAGGAGT-3' (SEQ ID NO: 13) which is the non-coding strand sequence which corresponds to coding strand sequence encoding amino acids 335-340 of *Cry j* I (ThrProGlnLeuThrLys, amino acids 335 through 340 of SEQ ID NO: 1). CP-18 has the sequence 5'-TAGCAACTCCAGTCGAAGT-3' (SEQ ID NO: 15) which is the non-coding strand sequence which substantially corresponds to coding strand sequence encoding amino acids 181 through 186 of *Cry j* I (ThrSerThrGlyValThr, amino acids 181 through 186 of SEQ ID NO: 1) except that the fourth nucleotide of CP-18 (SEQ ID NO: 15) was synthesized as a C rather than

the correct nucleotide, T. CP-19 which has the sequence 5'-TAGCTCTCATTTGGTGC-3' (SEQ ID NO: 16) is the non-coding strand sequence which corresponds to coding strand sequence encoding amino acids 270 through 275 of *Cry j* I (AlaProAsnGluSerTyr, amino acids 270 through 275 of SEQ ID NO: 1). CP-20 has the sequence 5'- TATGCAATTGGTGGGAGT-3' (SEQ ID NO: 17) which is the coding strand sequence for amino acids 251-256 of *Cry j* I (TyrAlaIleGlyGlySer, amino acids 251 through 256 of SEQ ID NO: 1). The sequenced DNA was found to have the sequence shown in Figs. 4a and 4b (SEQ ID NO: 1). This is a composite sequence from the two overlapping clones JC 71.6 and pUC19J91A. The complete cDNA sequence for *Cry j* I is composed of 1312 nucleotides, including 66 nucleotides of 5' untranslated sequence, an open reading frame starting with the codon for an initiating methionine, of 1122 nucleotides, and a 3' untranslated region. There is a consensus polyadenylation signal sequence in the 3' untranslated region 25 nucleotides 5' to the poly A tail. The position of the initiating methionine is confirmed by the presence of in- frame upstream stop codons and by 78% homology with the plant consensus sequence that encompasses the initiating methionine (AAAAAUGGA (bases 62 through 70 of SEQ ID NO: 1) found in *Cry j* I compared with the AACAAUGGC consensus sequence for plants, Lutcke et al. (1987) *EMBO J.* 6: 43-48). The open reading frame encodes a protein of 374 amino acids of which the first 21 amino acids comprise a leader sequence that is cleaved from the mature protein. The amino terminus of the mature protein was identified by comparison with the published NH_2-terminal sequence (Taniai et al. (1988) *supra*) and with sequence determined by direct amino acid analysis of purified native *Cry j* I. The deduced amino acid sequence of the mature protein, comprised of 353 amino acids has complete sequence identity with the published protein sequence for *Cry j* I (Taniai et al. *supra*), including the first twenty amino acids for the NH_2-terminal and sixteen contiguous internal amino acids. The mature protein also contains five potential N-linked glycosylation sites corresponding to the consensus sequence N-X-S/T.

Example 4

Extraction of RNA from Japanese Cedar Pollen Collected in Japan

Fresh pollen collected from a pool of *Cryptomeria japonica* (Japanese cedar) trees in Japan was frozen immediately on dry ice. RNA was prepared from 500 mg of the pollen, essentially as described by Frankis and Mascarenhas *Ann. Bot.* 45:595-599. The samples were ground by mortar and pestle on dry ice and suspended in 5 ml of 50 mM Tris pH 9.0 with 0.2 M NaCl, 1 mM EDTA, 1% SDS that had been treated overnight with 0.1% DEPC. After five extractions with phenol/chloroform/isoamyl alcohol (mixed at 25:24:1), the RNA was precipitated from the aqueous phase with 0.1 volume 3 M sodium acetate and 2 volumes ethanol. The pellets were recovered by centrifugation, resuspended in dH_20 and heated to 65°C for 5 minutes. Two ml of 4 M lithium chloride were added to the RNA preparations and they were incubated overnight at 9°C. The RNA pellets were recovered by centrifugation, resuspended in 1 ml dH_20, and again precipitated with 3 M sodium acetate and ethanol overnight. The final pellets were resuspended in 100 µl dH_20 and stored at -80°C.

Double stranded cDNA was synthesized from 8 µg pollen RNA using the cDNA Synthesis Systems kit (BRL) with oligo dT priming according to the method of Gubler and Hoffman (1983) *Gene* 25:263-269. Polymerase chain reactions (PCR) were carried out using the GeneAmp DNA Amplification kit (Perkin Elmer Cetus) whereby 10 µl 10x buffer containing dNTPs was mixed with 100 pmol each of a sense oligonucleotide and an anti-sense oligonucleotide, (10 µl of a 400 µl double stranded cDNA reaction mix), 0.5 µl Amplitaq DNA polymerase, and distilled water to 100 µl.

The samples were amplified with a programmable thermal controller from MJ Research, Inc. (Cambridge, MA). The first 5 rounds of amplification consisted of denaturation at 94°C for 1 minute, annealing of primers to the template at 45°C for 1 minute, and chain elongation at 72°C for 1 minute. The final 20 rounds of amplification consisted of denaturation as above, annealing at 55°C for 1 minute, and elongation as above.

Seven different *Cry j* I primer pairs were used to amplify the double stranded cDNA as follows: CP-9 (SEQ. ID #8) and CP-17 (SEQ. ID #14), CP-10 (SEQ. ID #9) and CP-17 (SEQ. ID #14), CP-10 (SEQ. ID #9) and CP-16 (SEQ. ID #13), CP-10 (SEQ. ID #9) and CP-19 (SEQ. ID #16), CP-10 (SEQ. ID #9) and CP-18 (SEQ. ID #15), CP-13 (SEQ. ID #10) and CP-17 (SEQ. ID #14), and CP-13 (SEQ. ID #10) and CP-19 (SEQ. ID #16). CP-17 (SEQ. ID #14) has the sequence 5'-

CCTGCAGAAGCTTCATCAACAACGTTTAGA-3' and corresponds to non-coding strand sequence that corresponds to coding strand sequence encoding amino acids SKRC* (amino acids 350-353 and the stop codon of SEQ. ID #1). The nucleotide sequence 5'-CCTGCAGAAGCTT-3' (bases 1 through 13 of SEQ. ID # 14) represents *Pst* I and *Hin* dIII restriction sites added for cloning purposes. The nucleotide sequence 5'-TCA-3' (bases 13 through 15 of SEQ. ID # 14) correspond to the non-coding strand sequence of a stop codon. All of the amplifications yielded products of the expected size when viewed on ethidium bromide (EtBr)-stained agarose gels. Two of these primer pairs were used in amplifications whose products were cloned into pUC19 for full-length sequencing. The PCR reaction with CP-10 (SEQ. ID #9) and CP-16 (SEQ. ID #13) on the double stranded cDNA yielded a band of approximately 1.1 kb, and was called JC130. A separate first strand cDNA reaction was done with 8 µg pollen RNA as described above and amplified with oligonucleotide primers CP-10 (SEQ. ID #9) and CP-17 (SEQ. ID #14). This amplification yielded a full-length cDNA, named JC135, from the amino terminus of the mature protein to the stop codon.

Amplified DNA was recovered by sequential chloroform, phenol, and chloroform extractions, followed by precipitation at -20°C with 0.5 volumes of 7.5 ammonium acetate and 1.5 volumes of isopropanol. After precipitation and washing with 70% ethanol, the DNA was blunted with T4 polymerase followed by digestion with *Eco* RI, in the case of JC130 , or simultaneously digested with *Eco* RI and *Pst* I, in the case of JC135, in a 15 µl reaction and electrophoresed through a preparative 1% SeaPlaque low melt gel (FMC). Appropriate sized DNA bands were visualized by EtBr staining, excised, and ligated into appropriately digested pUC19 for dideoxy DNA sequencing by the dideoxy chain termination method (Sanger et al. (1977) *Proc. Natl. Acad. Sci. USA* 74:5463-5476) using a commercially available sequencing kit (Sequenase kit, U.S. Biochemicals, Cleveland, OH).

Both strands were sequenced using M13 forward and reverse primers (N.E. Biolabs, Beverly, MA) and internal sequencing primers CP-13 (SEQ. ID #10), CP-15 (SEQ. ID #12), CP-16 (SEQ ID #13), CP-18 (SEQ. ID #15), CP-19 (SEQ. ID #16) and CP-20 (SEQ. ID #17). Two clones from amplification JC130 (JC130a and JC130b) and one clone from amplification JC135/(JC135g) were found to be *Cry j* I clones upon sequencing. The nucleotide and deduced amino acid sequences of clones JC130a and JC135g were identical to previously known *Cry j* I sequence

(SEQ. ID #1). Clone JC130b was found to contain a single nucleotide difference from the previously known *Cry j* I sequence (SEQ. ID #1). Clone JC130b had a T at nucleotide position 306 of Seq. ID #1 rather than the previously described C. This nucleotide change results in a predicted amino acid change from a Tyr to a His at amino acid 60 of the mature *Cry j* I protein. This polymorphism has not yet been confirmed in an independently-derived PCR clone or by direct amino acid sequencing. However, such polymorphisms in primary nucleotide and amino acid sequences are expected.

Example 5

Expression of *Cry j* I

Expression of *Cry j* I was performed as follows. Ten µg of pUC19JC91a was digested with *Xba* I, precipitated, then blunted with T4 polymerase. *Bam* HI linkers (N.E. Biolabs, Beverly, MA) were blunt-end ligated to pUC19JC91a overnight and excess linkers were removed by filtration through a NACS ion exchange minicolumn (BRL, Gaithersburg, MD). The linkered cDNA was then digested simultaneously with *Eco* RI and *Bam* HI. The *Cry j* I insert (extending from the nucleotides encoding the amino terminus of the mature protein through the stop codon) was isolated by electrophoresis of this digest through a 1% SeaPlaque low melt agarose gel. The insert was then ligated into the appropriately digested expression vector pET-11d (Novagen, Madison, WI; Jameel et al. (1990) *J. Virol.* 64:3963-3966) modified to contain a sequence encoding 6 histidines (His 6) immediately 3' of the ATG initiation codon followed by a unique *Eco* RI endonuclease restriction site. A second *Eco* RI endonuclease restriction site in the vector, along with neighboring *Cla* I and *Hind* III endonuclease restriction sites, had previously been removed by digestion with *Eco* RI and *Hind* III, blunted and religated. The histidine (His_6) sequence was added for affinity purification of the recombinant protein (*Cry j* I) on a Ni^{2+} chelating column (Hochuli et al. (1987) *J. Chromatog.* 411:177-184; Hochuli et al. (1988) *Bio/Tech.* 6:1321-1325.). A recombinant clone was used to transform *Escherichia coli* strain BL21-DE3 which harbors a plasmid that has an isopropyl-ß-D-thiogalactopyranoside (IPTG)-inducible promoter preceding the gene encoding T7 polymerase. Induction with IPTG leads to high levels of T7 polymerase expression, which is necessary for expression of the

recombinant protein in pET-11d, which has a T7 promoter. Clone pET-11dΔ HRhis6JC91a.d was confirmed by dideoxy sequencing (Sanger et al. *Supra*) with CP-14 (SEQ. ID #11) to be a *Cry j* I clone in the correct reading frame for expression.

Expression of the recombinant protein was confirmed in an initial small culture (50 ml). An overnight culture of clone pET-11dΔHRhis6JC91a.d was used to inoculate 50 ml of media (Brain Heart Infusion Media, Difco) containing ampicillin (200 µg/ml), grown to an A_{600} = 1.0 and then induced with IPTG (1 mM, final concentration) for 2 hrs. One ml aliquots of the bacteria were collected before and after induction, pelleted by centrifugation, and crude cell lysates prepared by boiling the pellets for 5 minutes in 50 mM Tris HCl, pH 6.8, 2 mM EDTA, 1% SDS, 1% ß-mercaptoethanol, 10% glycerol, 0.25% bromophenol blue (Studier et al., (1990) *Methods in Enzymology* 185:60-89). Recombinant protein expression was visualized as a band with the predicted molecular weight of approximately 38 kDa on a Coomassie blue-stained SDS-PAGE gel, according to the method in Sambrook et al., *supra*, on which 40 µl of the crude lysate was loaded. A negative control consisted of crude lysates from uninduced bacteria containing the plasmid with *Cry j* I and an induced lysate from bacteria carrying no plasmid.

The pET-11dΔ HRhis6JC91a.d clone was then grown on a large scale for recombinant protein expression and purification. A 2 ml culture bacteria containing the recombinant plasmid was grown for 8 hr, then streaked onto solid media (e.g. 6 petri plates (100 x 15 mm) with 1.5% agarose in LB medium (Gibco-BRL, Gaithersburg, MD) containing 200 µg/ml ampicillin), grown to confluence overnight, then scraped into 9 L of liquid media (Brain Heart Infusion media, Difco) containing ampicillin (200 µg/ml). The culture was grown until the A_{600} is 1.0, IPTG added (1 mM final concentration), and the culture grown for an additional 2 hours.

Bacteria was recovered by centrifugation (7,930 x g, 10 min), and lysed in 90 ml of 6M Guanidine-HCl, 0.1M Na_2HPO_4, pH 8.0 for 1 hour with vigorous shaking. Insoluble material was removed by centrifugation (11,000 x g, 10 min, 4° C). The pH of the lysate was adjusted to pH 8.0, and the lysate applied to an 80 ml Nickel NTA agarose column (Qiagen) that had been equilibrated with 6 M Guanidine HCl, 100 mM Na_2HPO_4, pH 8.0. The column was sequentially washed with 6 M Guanidine HCl, 100 mM Na_2HPO_4, 10 mM Tris-HCl, pH 8.0, then 8 M

urea, 100 mM Na_2HPO_4, pH 8.0, and finally 8 M urea, 100 mM sodium acetate, 10 mM Tris-HCl, pH 6.3. The column was washed with each buffer until the flow through has an $A_{280} \leq 0.05$.

The recombinant protein, *Cry j* I, was eluted with 8 M urea, 100 mM sodium acetate, 10 mM Tris-HCl, pH 4.5, and collected in 10 ml aliquots. The protein concentration of each fraction was determined by A_{280} and the peak fractions pooled. An aliquot of the collected recombinant protein was analyzed on SDS-PAGE according to the method in Sambrook et al., *supra*.

The first 9 L prep, JCpET-1, yielded 30 mg of *Cry j* I with approximately 78% purity, as determined by densitometry (Shimadzu Flying Spot Scanner, Shimadzu Scientific Instruments, Inc., Braintree, MA) of the Coomassie-blue stained SDS-PAGE gel. A second 9 L prep prepared the same way, JCpET-2, yielded 41 mg of *Cry j* I with approximately 77% purity.

Example 6

Japanese Cedar Pollen Allergic Patient T Cell Studies with *Cry j* I - the Primary Cedar Pollen Antigen.

T Cell Responses to Cedar Pollen Antigen Peptides

Peripheral blood mononuclear cells (PBMC) were purified by lymphocyte separation medium (LSM) centrifugation of 60 ml of heparinized blood from Japanese cedar pollen-allergic patients who exhibited clinical symptoms of seasonal rhinitis and were MAST and/or skin test positive for Japanese cedar pollen. Long term T cell lines were established by stimulation of 2 X 10^6 PBL/ml in bulk cultures of complete medium (RPMI-1640, 2 mM L-glutamine, 100 U/ml penicillin/streptomycin, $5x10^{-5}$M 2-mercaptoethanol, and 10 mM HEPES supplemented with 5% heat inactivated human AB serum) with 20 µg/ml of partially purified native *Cry j* I (75% purity containing three bands similar to the three bands in Fig. 2) for 7 days at 37°C in a humidified 5% CO_2 incubator to select for *Cry j* I reactive T cells. This amount of priming antigen was determined to be optimal for the activation of T cells from most cedar pollen allergic patients. Viable cells were purified by LSM centrifugation and cultured in complete medium supplemented with 5 units

recombinant human IL-2/ml and 5 units recombinant human IL-4/ml for up to three weeks until the cells no longer responded to lymphokines and were considered "rested". The ability of the T cells to proliferate to recombinant *Cry j* I (r*Cry j* I), purified native *Cry j* I, or recombinant *Amb a* I.1 (r*Amb a*I.1) was then assessed. For assay, 2 X 10^4 rested cells were restimulated in the presence of 4 X 10^4 autologous Epstein-Barr virus (EBV)-transformed B cells (prepared as described below) (gamma-irradiated with 25,000 RADS) with 2-50 µg/ml of r*Cry j* I, purified native *Cry j* I or r*Amb a* I.1, in a volume of 200 µl complete medium in duplicate or triplicate wells in 96-well round bottom plates for 2-4 days. The optimal incubation was found to be 3 days. Each well then received 1 µCi tritiated thymidine for 16-20 hours. The counts incorporated were collected onto glass fiber filter mats and processed for liquid scintillation counting. Fig. 12 shows the effect of varying antigen dose in assays with recombinant *Cry j* I, purified native *Cry j* I, and recombinant *Amb a* I.1. The results shown in Fig. 12 demonstrate that patient #999 responds well to recombinant *Cry j* I, and purified native *Cry j* I, but not to recombinant *Amb a* I.1. This indicates that *Cry j* I T cell epitopes are recognized by T cells from this particular allergic patient and that r*Cry j* I contains such T cell epitopes.

Preparation of (EBV)-transformed B Cells for Use as Antigen Presenting Cells

Autologous EBV-transformed cell lines were γ-irradiated with 25,000 Rad and used as antigen presenting cells in secondary proliferation assays and secondary bulk stimulations. These cell lines were also used as a control in the immunofluorescence flow cytometry analysis. These EBV-transformed cell lines were made by incubating 5 X 10^6 PBL with 1 ml of B-59/8 Marmoset cell line (ATCC CRL1612, American Type Culture Collection, Rockville, MD) conditioned medium in the presence of 1 µg/ml phorbol 12-myristate 13-acetate (PMA) at 37°C for 60 minutes in 12 X 75 mm polypropylene round-bottom Falcon snap cap tubes (Becton Dickinson LAbware, Lincoln Park, NJ). These cells were then diluted to 1.25 X 10^6 cells/ml in RPMI-1640 as described above except supplemented with 10% heat-inactivated fetal bovine serum and cultured in 200 µl aliquots in flat bottom culture plates until visible colonies were detected. They were then transferred to larger

wells until the cell lines were established.

Example 7

Cry j I as the Major Cedar Pollen Allergen

To examine the importance of *Cry j* I, reported as the major allergen of Japanese cedar pollen, both direct and competition ELISA assays were performed. For the direct ELISA assays, wells were coated with either soluble pollen extract (SPE) for Japanese cedar pollen or purified native *Cry j* I (assayed at 90% purity by protein sequencing) and human IgE antibody binding to these antigens was analyzed. Pooled human plasma, consisting of an equal volume of plasma from 15 patients with a Japanese cedar pollen MAST score of 2.5 or greater, and two individual patient plasma samples were compared in this assay. Fig. 5 shows the results of the binding reactivity with these two antigens. The overall pattern of binding is very similar whether the coating antigen is SPE (Fig. 5a) or purified native *Cry j* I (Fig. 5b).

In the competition assay, ELISA wells were coated with Japanese cedar pollen SPE and then allergic patient IgE binding was measured in the presence of competing purified native *Cry j* I in solution. The source of allergic IgE in these assays was either the pool of plasma from 15 patients (denoted PHP) or seven individual plasma samples from patients with a Japanese cedar MAST score of 2.5 or greater. The competition assay using the pooled human plasma samples compares the competitive binding capacity of purified native *Cry j* I to Japanese cedar pollen SPE and an irrelevant allergen source, rye grass SPE. Fig. 6 shows the graphed results of the competition ELISA with pooled human plasma. The concentration of protein present in the Japanese cedar pollen SPE is approximately 170 times greater at each competing point than is the purified native *Cry j* I. From this analysis it is clear that the purified native *Cry j* I competes very well for IgE binding to the whole range of proteins present in the Japanese cedar pollen soluble pollen extract. This implies that most of the anti-*Cry j* I IgE reactivity is directed against purified native *Cry j* I. The negative control shows no specific competitive activity and the competing SPE in solution can completely remove binding to the coated wells. This

assay was repeated with individual patients as a measure of the range of the IgE response within the allergic population. Fig. 7 shows this result where the competition of binding to SPE was performed with purified native *Cry j* I . The results demonstrate that although the patients show different dose response to Japanese cedar pollen SPE, each of the seven patients' IgE binding to Japanese cedar pollen SPE could be competed with purified native *Cry j* I. The implications of these data are that for each patient the IgE reactivity directed against *Cry j* I is predominant but that there is variation in this reactivity between patients. The overall conclusion is that these data support the previous findings (Yasueda et al., (1988) *supra*) that *Cry j* I is the major allergen of Japanese cedar pollen.

The reactivity of IgE from cedar pollen allergic patients to the pollen proteins is dramatically reduced when these proteins are denatured. One method of analyzing this property is through direct binding ELISA where the coating antigen is the Japanese cedar pollen SPE or denatured Japanese cedar pollen SPE which has been denatured by boiling in the presence of a reducing agent DTT. This is then examined with allergic patient plasma for IgE binding reactivity. Fig. 8a, shows the direct binding assay to the SPE with seven individual plasma samples. In Fig. 8b, the binding results with the denatured SPE demonstrates the marked decrease in reactivity following this treatment. To determine the extent of *Cry j* I binding to the ELISA wells, *Cry j* I was detected with a rabbit polyclonal antisera against the *Amb a* I & II protein family. These proteins have high sequence identity (46%) with *Cry j* I and this antisera can be used as a cross reactive antibody detection system. In conclusion, these data demonstrate a marked loss in IgE reactivity following denaturation of the SPE.

Example 8

IgE Reactivity and Histamine Release Analysis

The recombinant *Cry j* I protein (r*Cry j* I), expressed in bacteria and then purified (as described in Example 5), has been examined for IgE reactivity. The first method applied to this examination was direct ELISA where wells were coated with the recombinant *Cry j* I and IgE binding was assayed on individual patients. Fig. 9 is the graphic representation of this direct ELISA. The only positive signals on this

data set are from the two control antisera rabbit polyclonal anti-*Amb a* I & II (Rabbit anti-*Amb a* I & II) and CBF2, a monoclonal antibody raised against *Amb a* I that cross reacts with *Cry j* I. By this method all patients tested showed no IgE reactivity with the recombinant *Cry j* I.

Another method of analysis that was applied to the examination of IgE reactivity to the recombinant *Cry j* I was a capture ELISA. This analysis relies on the use of a defined anitibody, in this case CBF2 to bind the antigen and allow for the binding of antibodies to other epitope sites. The format of this capture ELISA is 1) wells are coated with MAb CBF2, 2) antigen or PBS (as one type of negative control) is added and captured by specific interaction with the coated MAb, 3) either the control antibody anti-*Amb a* I & II (Fig. 10b) or human allergic plasma (Fig. 10a) is added as the detecting antibody, and 4) detection of antibody binding is assayed. Figs. 10a and 10b are the graphed results of these assays. For the IgE analysis, the pooled human plasma (15 patients) was used. The conclusion from these results is that there is no indication of any specific binding of human allergic IgE to r*Cry j* I by this method of analysis. However, the capture of r*Cry j* I works as evidenced by the control antibody binding curve, shown in Fig. 10b. The lack of IgE binding to r*Cry j* I may be due to absence of carbohydrate or any other post-translational modification and/or that the majority of IgE cannot react with denatured *Cry j* I. RAST, competition ELISA and Western blotting data also demonstrates no specific IgE reactivity to the r*Cry j* I (data not shown).

A histamine release assay was performed on one Japanese cedar pollen allergic patient using Japanese cedar pollen SPE, purified native *Cry j* I and r*Cry j* I as the added antigens. This assay is a measure of IgE reactivity through human basophil mediator release. The results of this assay, shown in Fig. 11, demonstrate strong histamine release with both purified native *Cry j* I and the Japanese cedar pollen SPE over a wide concentration range. The only point where there is any measurable histamine release with the *Cry j* I is at the highest concentration, 50 µg/ml. Two possible explanations for this release by the r*Cry j* I are: 1) specific reactivity with a very low proportion of the anti-*Cry j* I IgE capable of recognizing the recombinant form of *Cry j* I, or 2) non-specific release caused by low abundance of bacterial contaminants observed only at the highest antigen concentration. Thus far, this result has only been shown in a single patient. In addition, the data shown are from single data points at each protein concentration.

It may be possible to use this recombinantly expressed *Cry j* I protein for immunotherapy as *E. coli* expressed material has T cell reactivity (Example 6), but does not appear to bind IgE from *Cryptomeria japonica* atopes nor cause histamine release from the mast cells and basophils of such atopes *in vitro*. Expression of r*Cry j* I which is capable of binding IgE could be achieved in yeast, insect (baculovirus) or mammalian cells (e.g. CHO, human and mouse). A r*Cry j* I capable of actively binding IgE may be important for the use of recombinant material for diagnostic purposes.

Although the invention has been described with reference to its preferred embodiments, other embodiments, can achieve the same results. Variations and modifications to the present invention will be obvious to those skilled in the art and it is intended to cover in the appended claims all such modification and equivalents and follow in the true spirit and scope of this invention.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Griffith, Irwin J.
Pollock, Joanne, Bond Julian

(ii) TITLE OF INVENTION: Allergenic Proteins And Peptides From Japanese Cedar Pollen

(iii) NUMBER OF SEQUENCES: 25

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: ImmuLogic Pharmaceutical Corporation
(B) STREET: One Kendall Square, Building 600
(C) CITY: Boston
(D) STATE: MA
(E) COUNTRY: USA
(F) ZIP: 02139

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Stacey L. Channing
(B) REGISTRATION NUMBER: 31,095
(C) REFERENCE/DOCKET NUMBER: IPC-025CPCT

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (617) 494-0060
(B) TELEFAX: (617) 494-4964

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1337 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Crytpomeria japonica

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 66..1187

(ix) FEATURE:
(A) NAME/KEY: mat_peptide
(B) LOCATION: 129..1187

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
AGTCAATCTG CTCATAATCA TAGCATAGCC GTATAGAAAG AAATTCTACA CTCTGCTACC    60

AAAAA ATG GAT TCC CCT TGC TTA GTA GCA TTA CTG GTT TTC TCT TTT        107
      Met Asp Ser Pro Cys Leu Val Ala Leu Leu Val Phe Ser Phe
      -21 -20                 -15                 -10

GTA ATT GGA TCT TGC TTT TCT GAT AAT CCC ATA GAC AGC TGC TGG AGA      155
Val Ile Gly Ser Cys Phe Ser Asp Asn Pro Ile Asp Ser Cys Trp Arg
         -5                  1                   5

GGA GAC TCA AAC TGG GCC CAA AAT AGA ATG AAG CTC GCA GAT TGT GCA      203
Gly Asp Ser Asn Trp Ala Gln Asn Arg Met Lys Leu Ala Asp Cys Ala
 10                  15                  20                  25

GTG GGC TTC GGA AGC TCC ACC ATG GGA GGC AAG GGA GGA GAT CTT TAT      251
Val Gly Phe Gly Ser Ser Thr Met Gly Gly Lys Gly Gly Asp Leu Tyr
                 30                  35                  40

ACG GTC ACG AAC TCA GAT GAC GAC CCT GTG AAT CCT GCA CCA GGA ACT      299
Thr Val Thr Asn Ser Asp Asp Asp Pro Val Asn Pro Ala Pro Gly Thr
             45                  50                  55

CTG CGC TAT GGA GCA ACC CGA GAT AGG CCC CTG TGG ATA ATT TTC AGT      347
Leu Arg Tyr Gly Ala Thr Arg Asp Arg Pro Leu Trp Ile Ile Phe Ser
         60                  65                  70

GGG AAT ATG AAT ATA AAG CTC AAA ATG CCT ATG TAC ATT GCT GGG TAT      395
Gly Asn Met Asn Ile Lys Leu Lys Met Pro Met Tyr Ile Ala Gly Tyr
     75                  80                  85

AAG ACT TTT GAT GGC AGG GGA GCA CAA GTT TAT ATT GGC AAT GGC GGT      443
Lys Thr Phe Asp Gly Arg Gly Ala Gln Val Tyr Ile Gly Asn Gly Gly
```

```
 90                      95                      100                     105

CCC TGT GTG TTT ATC AAG AGA GTT AGC AAT GTT ATC ATA CAC GGT TTG   491
Pro Cys Val Phe Ile Lys Arg Val Ser Asn Val Ile Ile His Gly Leu
                110                 115                 120

TAT CTG TAC GGC TGT AGT ACT AGT GTT TTG GGG AAT GTT TTG ATA AAC   539
Tyr Leu Tyr Gly Cys Ser Thr Ser Val Leu Gly Asn Val Leu Ile Asn
            125                 130                 135

GAG AGT TTT GGG GTG GAG CCT GTT CAT CCT CAG GAT GGC GAT GCT CTT   587
Glu Ser Phe Gly Val Glu Pro Val His Pro Gln Asp Gly Asp Ala Leu
        140                 145                 150

ACT CTG CGC ACT GCT ACA AAT ATT TGG ATT GAT CAT AAT TCT TTC TCC   635
Thr Leu Arg Thr Ala Thr Asn Ile Trp Ile Asp His Asn Ser Phe Ser
    155                 160                 165

AAT TCT TCT GAT GGT CTG GTC GAT GTC ACT CTT ACT TCG ACT GGA GTT   683
Asn Ser Ser Asp Gly Leu Val Asp Val Thr Leu Thr Ser Thr Gly Val
170                 175                 180                 185

ACT ATT TCA AAC AAT CTT TTT TTC AAC CAT CAT AAA GTG ATG TTG TTA   731
Thr Ile Ser Asn Asn Leu Phe Phe Asn His His Lys Val Met Leu Leu
                190                 195                 200

GGG CAT GAT GAT GCA TAT AGT GAT GAC AAA TCC ATG AAG GTG ACA GTG   779
Gly His Asp Asp Ala Tyr Ser Asp Asp Lys Ser Met Lys Val Thr Val
            205                 210                 215

GCG TTC AAT CAA TTT GGA CCT AAC TGT GGA CAA AGA ATG CCC AGG GCA   827
Ala Phe Asn Gln Phe Gly Pro Asn Cys Gly Gln Arg Met Pro Arg Ala
        220                 225                 230

CGA TAT GGA CTT GTA CAT GTT GCA AAC AAT AAT TAT GAC CCA TGG ACT   875
Arg Tyr Gly Leu Val His Val Ala Asn Asn Asn Tyr Asp Pro Trp Thr
    235                 240                 245

ATA TAT GCA ATT GGT GGG AGT TCA AAT CCA ACC ATT CTA AGT GAA GGG   923
Ile Tyr Ala Ile Gly Gly Ser Ser Asn Pro Thr Ile Leu Ser Glu Gly
250                 255                 260                 265

AAT AGT TTC ACT GCA CCA AAT GAG AGC TAC AAG AAG CAA GTA ACC ATA   971
Asn Ser Phe Thr Ala Pro Asn Glu Ser Tyr Lys Lys Gln Val Thr Ile
                270                 275                 280

CGT ATT GGA TGC AAA ACA TCA TCA TCT TGT TCA AAT TGG GTG TGG CAA  1019
Arg Ile Gly Cys Lys Thr Ser Ser Ser Cys Ser Asn Trp Val Trp Gln
            285                 290                 295

TCT ACA CAA GAT GTT TTT TAT AAT GGA GCT TAT TTT GTA TCA TCA GGG  1067
Ser Thr Gln Asp Val Phe Tyr Asn Gly Ala Tyr Phe Val Ser Ser Gly
        300                 305                 310

AAA TAT GAA GGG GGT AAT ATA TAC ACA AAG AAA GAA GCT TTC AAT GTT  1115
Lys Tyr Glu Gly Gly Asn Ile Tyr Thr Lys Lys Glu Ala Phe Asn Val
    315                 320                 325

GAG AAT GGG AAT GCA ACT CCT CAA TTG ACA AAA AAT GCT GGG GTT TTA  1163
Glu Asn Gly Asn Ala Thr Pro Gln Leu Thr Lys Asn Ala Gly Val Leu
330                 335                 340                 345
```

ACA TGC TCT CTC TCT AAA CGT TGT TGATGATGCA TATATTCTAG CATGTTGTAC 1217
Thr Cys Ser Leu Ser Lys Arg Cys
350

TATCTAAATT AACATCAACA AGAAAAATATA TCATGATGTA TATTGTTGTA TTGATGTCAA 1277

AATAAAAATG TATCTTTTAC TATTAAAAAA AAAAATGATC GATCGGACGG TACCTCTAGA 1337

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 374 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Asp Ser Pro Cys Leu Val Ala Leu Leu Val Phe Ser Phe Val Ile
-21 -20 -15 -10

Gly Ser Cys Phe Ser Asp Asn Pro Ile Asp Ser Cys Trp Arg Gly Asp
-5 1 5 10

Ser Asn Trp Ala Gln Asn Arg Met Lys Leu Ala Asp Cys Ala Val Gly
15 20 25

Phe Gly Ser Ser Thr Met Gly Gly Lys Gly Gly Asp Leu Tyr Thr Val
30 35 40

Thr Asn Ser Asp Asp Asp Pro Val Asn Pro Ala Pro Gly Thr Leu Arg
45 50 55

Tyr Gly Ala Thr Arg Asp Arg Pro Leu Trp Ile Ile Phe Ser Gly Asn
60 65 70 75

Met Asn Ile Lys Leu Lys Met Pro Met Tyr Ile Ala Gly Tyr Lys Thr
80 85 90

Phe Asp Gly Arg Gly Ala Gln Val Tyr Ile Gly Asn Gly Gly Pro Cys
95 100 105

Val Phe Ile Lys Arg Val Ser Asn Val Ile Ile His Gly Leu Tyr Leu
110 115 120

Tyr Gly Cys Ser Thr Ser Val Leu Gly Asn Val Leu Ile Asn Glu Ser
125 130 135

Phe Gly Val Glu Pro Val His Pro Gln Asp Gly Asp Ala Leu Thr Leu
140 145 150 155

Arg Thr Ala Thr Asn Ile Trp Ile Asp His Asn Ser Phe Ser Asn Ser
160 165 170

Ser Asp Gly Leu Val Asp Val Thr Leu Thr Ser Thr Gly Val Thr Ile
175 180 185

```
Ser Asn Asn Leu Phe Phe Asn His His Lys Val Met Leu Leu Gly His
        190                 195                 200

Asp Asp Ala Tyr Ser Asp Asp Lys Ser Met Lys Val Thr Val Ala Phe
    205                 210                 215

Asn Gln Phe Gly Pro Asn Cys Gly Gln Arg Met Pro Arg Ala Arg Tyr
220                 225                 230                 235

Gly Leu Val His Val Ala Asn Asn Asn Tyr Asp Pro Trp Thr Ile Tyr
                240                 245                 250

Ala Ile Gly Gly Ser Ser Asn Pro Thr Ile Leu Ser Glu Gly Asn Ser
            255                 260                 265

Phe Thr Ala Pro Asn Glu Ser Tyr Lys Lys Gln Val Thr Ile Arg Ile
        270                 275                 280

Gly Cys Lys Thr Ser Ser Ser Cys Ser Asn Trp Val Trp Gln Ser Thr
    285                 290                 295

Gln Asp Val Phe Tyr Asn Gly Ala Tyr Phe Val Ser Ser Gly Lys Tyr
300                 305                 310                 315

Glu Gly Gly Asn Ile Tyr Thr Lys Lys Glu Ala Phe Asn Val Glu Asn
                320                 325                 330

Gly Asn Ala Thr Pro Gln Leu Thr Lys Asn Ala Gly Val Leu Thr Cys
            335                 340                 345

Ser Leu Ser Lys Arg Cys
        350
```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GAYAAYCCNA THGAYWS

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GGGAATTCAA YTGGGCNCAR AAYSG 25

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 15
(D) OTHER INFORMATION: /mod_base= i

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CTGCAGCCRT TYTCNACRTT RAA 23

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 6
(D) OTHER INFORMATION: /mod_base= i

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

TTCATNCKRT TYTGNGCCCA 20

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CCTGCAGCKR TTYTGNGCCC AARTT 25

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

ATGGATTCCC CTTGCTTA 18

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GGGAATTCGA TAATCCCATA GACAGC 26

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

ATGCCTATGT ACATTGC 17

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GCAATGTACA TAGGCAT

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

TCCAATTCTT CTGATGGT

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

TTTTGTCAAT TGAGGAGT

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

CCTGCAGAAG CTTCATCAAC AACGTTTAGA

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

TAGCAACTC AGTCGAAGT

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

TAGCTCTCAT TTGGTGC

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

TATGCAATTG GTGGGAGT

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: N-terminal

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Cryptomeria japonica

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 7
(D) OTHER INFORMATION: /note= "the amino acid at position 7 is Ser, Cys, Thr, or His"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Asp Asn Pro Ile Asp Ser Xaa Trp Arg Gly Asp Ser Asn Trp Ala Gln
1 5 10 15

Asn Arg Met Lys
20

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Cryptomeria japonica

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

Glu Ala Phe Asn Val Glu Asn Gly Asn Ala Thr Pro Gln Leu Thr Lys
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

GGGTCTAGAG GTACCGTCCG ATCGATCATT

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

GGGTCTAGAG GTACCGTCCG

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 13 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

AATGATCGAT GCT

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

GGAATTCTCT AGACTGCAGG T

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 35 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

GGAATTCTCT AGACTGCAGG TTTTTTTTTT TTTTT

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: N-terminal

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Juniperus sabinoides

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

Asp Asn Pro Ile Asp
1 5

What is claimed is:

1. A nucleic acid sequence coding for the Japanese cedar pollen allergen *Cry j* I, or at least one antigenic fragment thereof or the functional equivalent of said nucleic acid sequence.

2. The nucleic acid sequence of claim 1 wherein said nucleic acid sequence has the nucleotide sequence of bases 66 through 1187 of SEQ ID NO: 1.

3. The nucleic acid sequence of claim 1 wherein said nucleic acid sequence has the nucleotide sequence of bases 129 through 1187 of SEQ ID NO: 1.

4. A nucleic acid sequence of claim 1 wherein said nucleic acid sequence consists essentially of at least one fragment of the coding portion of the nucleic acid sequence of SEQ ID NO: 1.

5. An expression vector comprising a nucleic acid sequence coding for the Japanese cedar pollen allergen *Cry j* I, or at least one antigenic fragment thereof or the functional equivalent of said nucleic acid sequence.

6. The expression vector of claim 5 wherein said nucleic acid sequence has the nucleotide sequence of bases 66 through 1187 of SEQ ID NO: 1.

7. The expression vector of claim 5 wherein said nucleic acid sequence has the nucleotide sequence of bases 129 through 1187 of SEQ ID NO: 1

8. The expression vector of claim 5 wherein said nucleic acid sequence consists essentially of at least one fragment of the coding portion of the nucleic acid sequence of SEQ ID NO: 1.

9. A host cell transformed to express a protein or peptide encoded by the nucleic acid sequence of claim 1, 2, 3 or 4.

10. A host cell of claim 9 wherein said host cell is *E .coli.*

11. Purified Japanese cedar pollen allergen *Cry j* I or at least one antigenic fragment thereof produced in a host cell transformed with the nucleic acid sequence of claim 1 , 2 ,3 or 4.

12. Purified Japanese cedar pollen allergen of claim 11 wherein said Japanese cedar pollen allergen does not bind immunoglobulin E specific for Japanese cedar pollen or if binding of the Japanese cedar pollen allergen to said immunoglobulin E occurs, such binding does not result in histamine release from mast cells or basophils.

13. Purified Japanese cedar pollen allergen of claim 11 wherein said Japanese cedar pollen allergen binds immunoglobulin E to a substantially lesser extent than purifed native Japanese cedar pollen allergen binds said immunoglobulin E.

14. The purified Japanese cedar pollen allergen or antigenic fragment thereof of claim 11 wherein the host cell is *E. coli:*

15. A method of producing Japanese cedar pollen allergen *Cry j* I or at least one fragment thereof comprising the steps of:

a) culturing a host cell transformed with a DNA sequence encoding Japanese cedar pollen allergen *Cry j* I or fragment thereof in an appropriate medium to produce a mixture of cells and medium containing said Japanese cedar pollen allergen *Cry j* I or at least one fragment thereof; and

b) purifying said mixture to produce substantially pure Japanese cedar pollen allergen *Cry j* I, or at least one fragment thereof.

16. A protein preparation comprising Japanese cedar pollen allergen *Cry j* I, or at least one fragment thereof synthesized in a host cell transformed with a nucleic acid sequence encoding all or a portion of Japanese cedar pollen allergen *Cry j* I.

17. The protein preparation of claim 16 wherein said at least one fragment of *Cry j* I is an antigenic fragment.

18. A protein preparation comprising chemically synthesized Japanese cedar pollen allergen *Cry j* I or at least one fragment thereof.

19. The protein preparation of claim 16 or 18 wherein said *Cry j* I has the amino acid sequence of SEQ ID NO: 1.

20. An isolated antigenic fragment of an allergen from Japanese cedar pollen.

21. The antigenic fragment of claim 20 wherein said allergen from Japanese cedar pollen is *Cry j* I.

22. The antigenic fragment of claim 20 or 21 wherein said antigenic fragment comprises at least one T cell epitope.

23. The antigenic fragment of claim 22 wherein said antigenic fragment has minimal immunoglobulin E stimulating activity.

24. The antigenic fragment of claim 22 wherein said antigenic fragment does not bind immunoglobulin E specific for Japanese cedar pollen or if binding of the fragment to said immunoglobulin E occurs, such binding does not result in histamine

release from mast cells or basophils.

25. The antigenic fragment of claim 20 wherein said antigenic fragment binds immunoglobulin E to a substantially lesser extent than purified native Japanese cedar pollen allergen binds said immunoglobulin E.

26. The purified allergen or antigenic fragment of claim 11, 20, 21 or 22 wherein said purified allergen or said antigenic fragment is capable of modifying, in a Japanese cedar pollen-sensitive individual to which it is administered, the allergic response to Japanese cedar pollen.

27. The purified allergen or antigenic fragment of claim 26 wherein said purified allergen or said antigenic fragment is capable of modifying B-cell response of the individual to a Japanese cedar pollen allergen, T-cell response of the individual to a Japanese cedar pollen antigen, or both the B cell response and the T cell response of the individual to Japanese cedar pollen allergen.

28. A nucleic acid sequence coding for the isolated antigenic fragment of Japanese cedar pollen allergen of claim 20.

29. A modified Japanese cedar pollen allergen which, when administered to a Japanese cedar pollen-sensitive individual, reduces the allergic response of the individual to Japanese cedar pollen allergen.

30. The modified cedar pollen protein allergen of claim 29 wherein said modified Japanese cedar pollen allergen is a modified *Cry j* I protein.

31. At least one modified fragment of Japanese cedar pollen allergen which, when administered to a Japanese cedar pollen-sensitive individual, reduces the allergic response of the individual to Japanese cedar pollen allergen.

32. At least one modified fragment of claim 31 wherein said at least one modified fragment is a modified fragment of *Cry j* I protein.

33. An isolated protein allergen or antigenic fragment thereof that is immunologically related to *Cry j* I or fragment thereof.

34. The isolated protein allergen or antigenic fragment thereof of claim 33 wherein said protein allergen or antigenic fragment thereof binds to antibodies specific for *Cry j* I or a fragment thereof.

35. The isolated protein allergen or fragment thereof of claim 33 wherein said isolated protein allergen or antigenic fragment thereof is capable of stimulating T cells specific for *Cry j I* or a fragment thereof.

36. A therapeutic composition comprising purified Japanese cedar pollen

allergen *Cry j* I or at least one fragment thereof and a pharmaceutically acceptable carrier or diluent.

37. The therapeutic composition of claim 36 wherein *Cry j* I has the sequence of amino acids 1-353 of SEQ ID NO: 1.

38. A method of treating sensitivity to Japanese cedar pollen allergen or an allergen immulogically cross reactive with Japanese cedar pollen allergen in a mammal sensitive to said allergen, comprising administering to said mammal a therapeutically effective amount of said protein preparation of claim 16 or claim 18.

39. The protein preparation of claim 16 or claim 18 for use in therapy, e.g. in the treatment for sensitivity in an individual to Japanese cedar pollen allergen or an allergen cross reactive with Japanese cedar pollen allergen.

40. A method of detecting sensitivity in a mammal to a Japanese cedar pollen allergen comprising combining a blood sample obtained from said mammal with a purified Japanese cedar pollen allergen or antigenic fragment thereof produced in a host cell transformed with the nucleic acid sequence of claim 1 or chemically synthesized under conditions appropriate for binding of blood components with the protein or fragment thereof and determining the extent to which such binding occurs.

41. The method of claim 40 wherein the extent to which binding occurs is determined by assessing T cell function, T cell proliferation, B cell function, binding of the protein or fragment thereof to antibodies present in the blood or a combination thereof.

42. A method of detecting sensitivity of a mammal to Japanese cedar pollen allergen comprising administering to said mammal a sufficient quantity of the Japanese cedar pollen allergen *Cry j* I or at least one antigenic fragment thereof produced in a host cell transformed with the nucleic acid sequence of claim 1 or chemically synthesized to provoke an allergic response in said mammal and determining the occurrence of an allergic response in the individual to said Japanese cedar pollen allergen or antigenic fragment thereof.

43. A monoclonal antibody specifically reactive with a Japanese cedar pollen allergen, *Cry j* I, or at least one antigenic fragment thereof.

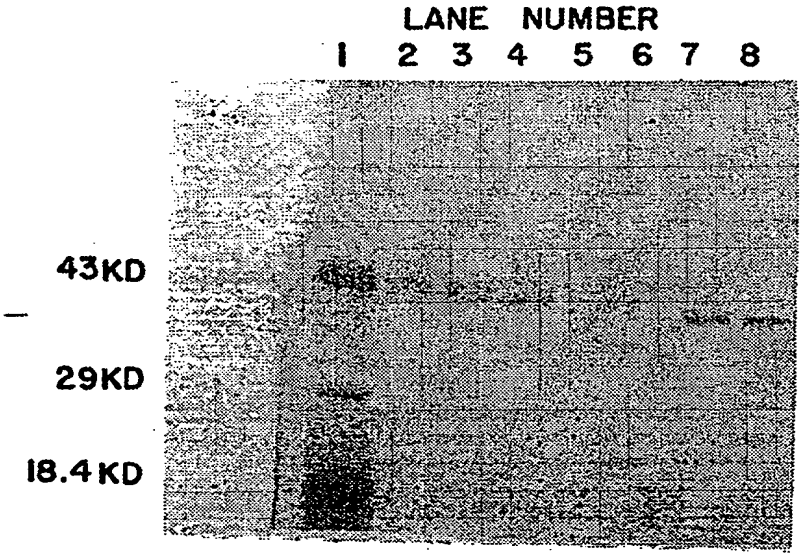


Fig. 1b

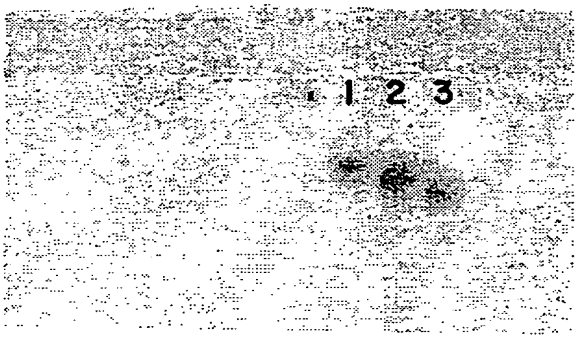


Fig. 2

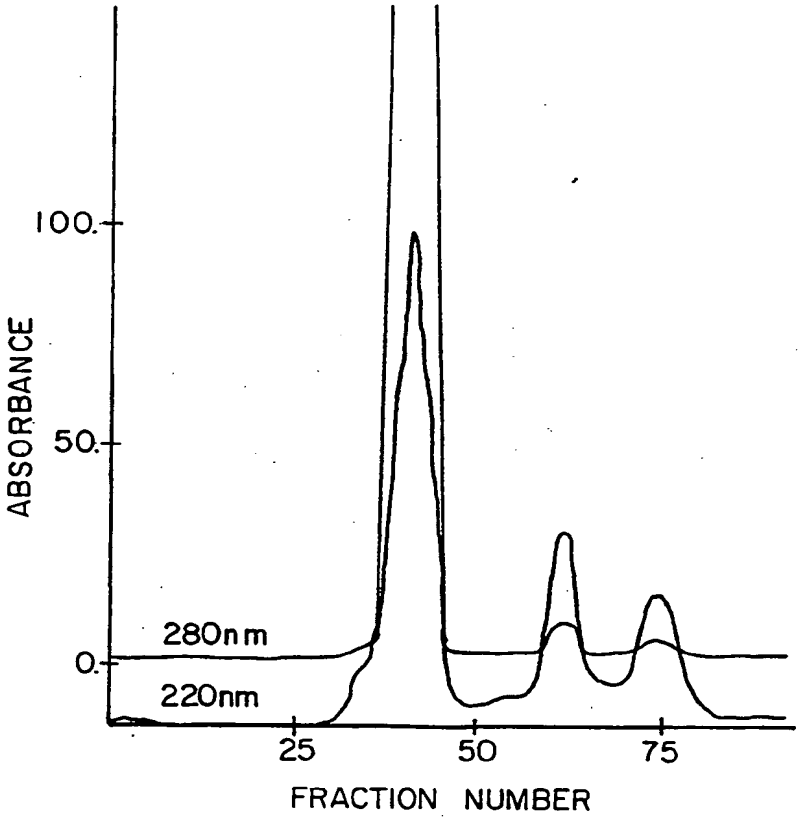

100.
50.
0.
ABSORBANCE
280nm
220nm
25
50
75
FRACTION NUMBER

Fig. 1a

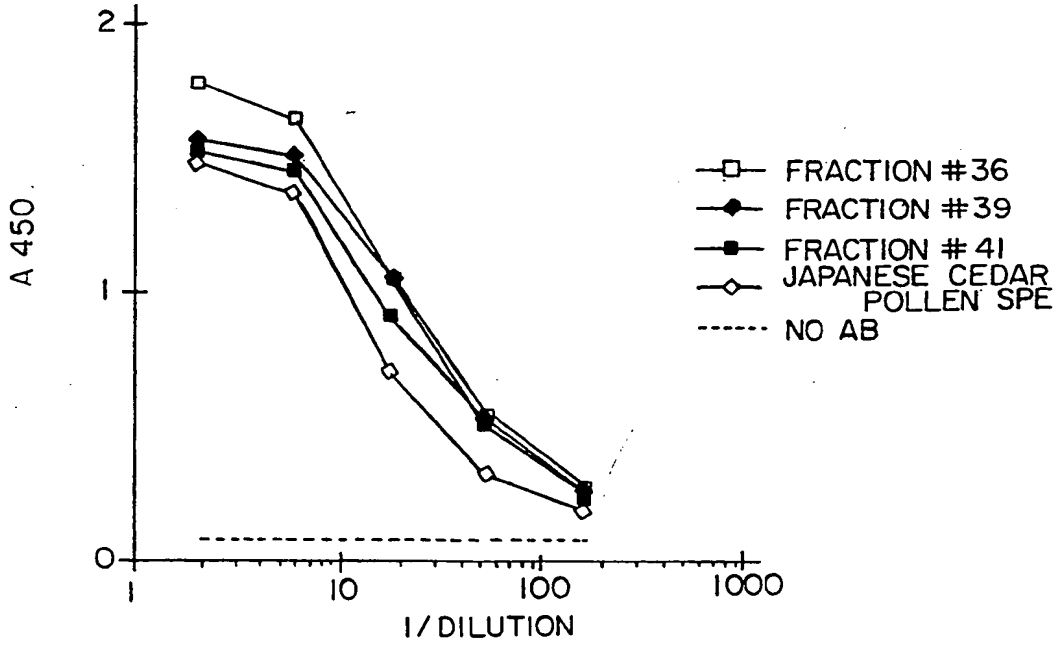

2
1
0
A 450
1
10
100
1000
1/DILUTION
FRACTION #36
FRACTION #39
FRACTION #41
JAPANESE CEDAR POLLEN SPE
NO AB

Fig. 3

```
5'-AGTCAATCTG CTCATAATCA TAGCATAGCC GTATAGAAAG AAATTCTACA CTCTGCTACC   60

AAAAA ATG GAT TCC CCT TGC TTA GTA GCA TTA CTG GTT TTC TCT TTT          107
      Met Asp Ser Pro Cys Leu Val Ala Leu Leu Val Phe Ser Phe
      -21 -20                 -15                 -10

GTA ATT GGA TCT TGC TTT TCT GAT AAT CCC ATA GAC AGC TGC TGG AGA        155
Val Ile Gly Ser Cys Phe Ser Asp Asn Pro Ile Asp Ser Cys Trp Arg
         -5                  1                   5

GGA GAC TCA AAC TGG GCC CAA AAT AGA ATG AAG CTC GCA GAT TGT GCA        203
Gly Asp Ser Asn Trp Ala Gln Asn Arg Met Lys Leu Ala Asp Cys Ala
 10                  15                  20                  25

GTG GGC TTC GGA AGC TCC ACC ATG GGA GGC AAG GGA GGA GAT CTT TAT        251
Val Gly Phe Gly Ser Ser Thr Met Gly Gly Lys Gly Gly Asp Leu Tyr
                 30                  35                  40

ACG GTC ACG AAC TCA GAT GAC GAC CCT GTG AAT CCT GCA CCA GGA ACT        299
Thr Val Thr Asn Ser Asp Asp Asp Pro Val Asn Pro Ala Pro Gly Thr
             45                  50                  55

CTG CGC TAT GGA GCA ACC CGA GAT AGG CCC CTG TGG ATA ATT TTC AGT        347
Leu Arg Tyr Gly Ala Thr Arg Asp Arg Pro Leu Trp Ile Ile Phe Ser
         60                  65                  70

GGG AAT ATG AAT ATA AAG CTC AAA ATG CCT ATG TAC ATT GCT GGG TAT        395
Gly Asn Met Asn Ile Lys Leu Lys Met Pro Met Tyr Ile Ala Gly Tyr
     75                  80                  85

AAG ACT TTT GAT GGC AGG GGA GCA CAA GTT TAT ATT GGC AAT GGC GGT        443
Lys Thr Phe Asp Gly Arg Gly Ala Gln Val Tyr Ile Gly Asn Gly Gly
 90                  95                  100                 105

CCC TGT GTG TTT ATC AAG AGA GTT AGC AAT GTT ATC ATA CAC GGT TTG        491
Pro Cys Val Phe Ile Lys Arg Val Ser Asn Val Ile Ile His Gly Leu
                110                 115                 120

TAT CTG TAC GGC TGT AGT ACT AGT GTT TTG GGG AAT GTT TTG ATA AAC        539
Tyr Leu Tyr Gly Cys Ser Thr Ser Val Leu Gly Asn Val Leu Ile Asn
            125                 130                 135

GAG AGT TTT GGG GTG GAG CCT GTT CAT CCT CAG GAT GGC GAT GCT CTT        587
Glu Ser Phe Gly Val Glu Pro Val His Pro Gln Asp Gly Asp Ala Leu
        140                 145                 150

ACT CTG CGC ACT GCT ACA AAT ATT TGG ATT GAT CAT AAT TCT TTC TCC        635
Thr Leu Arg Thr Ala Thr Asn Ile Trp Ile Asp His Asn Ser Phe Ser
    155                 160                 165
```

Fig. 4a

```
AAT TCT TCT GAT GGT CTG GTC GAT GTC ACT CTT ACT TCG ACT GGA GTT      683
Asn Ser Ser Asp Gly Leu Val Asp Val Thr Leu Thr Ser Thr Gly Val
170                 175                 180                 185

ACT ATT TCA AAC AAT CTT TTT TTC AAC CAT CAT AAA GTG ATG TTG TTA      731
Thr Ile Ser Asn Asn Leu Phe Phe Asn His His Lys Val Met Leu Leu
                190                 195                 200

GGG CAT GAT GAT GCA TAT AGT GAT GAC AAA TCC ATG AAG GTG ACA GTG      779
Gly His Asp Asp Ala Tyr Ser Asp Asp Lys Ser Met Lys Val Thr Val
            205                 210                 215

GCG TTC AAT CAA TTT GGA CCT AAC TGT GGA CAA AGA ATG CCC AGG GCA      827
Ala Phe Asn Gln Phe Gly Pro Asn Cys Gly Gln Arg Met Pro Arg Ala
        220                 225                 230

CGA TAT GGA CTT GTA CAT GTT GCA AAC AAT AAT TAT GAC CCA TGG ACT      875
Arg Tyr Gly Leu Val His Val Ala Asn Asn Asn Tyr Asp Pro Trp Thr
    235                 240                 245

ATA TAT GCA ATT GGT GGG AGT TCA AAT CCA ACC ATT CTA AGT GAA GGG      923
Ile Tyr Ala Ile Gly Gly Ser Ser Asn Pro Thr Ile Leu Ser Glu Gly
250                 255                 260                 265

AAT AGT TTC ACT GCA CCA AAT GAG AGC TAC AAG AAG CAA GTA ACC ATA      971
Asn Ser Phe Thr Ala Pro Asn Glu Ser Tyr Lys Lys Gln Val Thr Ile
                270                 275                 280

CGT ATT GGA TGC AAA ACA TCA TCA TCT TGT TCA AAT TGG GTG TGG CAA     1019
Arg Ile Gly Cys Lys Thr Ser Ser Ser Cys Ser Asn Trp Val Trp Gln
            285                 290                 295

TCT ACA CAA GAT GTT TTT TAT AAT GGA GCT TAT TTT GTA TCA TCA GGG     1067
Ser Thr Gln Asp Val Phe Tyr Asn Gly Ala Tyr Phe Val Ser Ser Gly
        300                 305                 310

AAA TAT GAA GGG GGT AAT ATA TAC ACA AAG AAA GAA GCT TTC AAT GTT     1115
Lys Tyr Glu Gly Gly Asn Ile Tyr Thr Lys Lys Glu Ala Phe Asn Val
    315                 320                 325

GAG AAT GGG AAT GCA ACT CCT CAA TTG ACA AAA AAT GCT GGG GTT TTA     1163
Glu Asn Gly Asn Ala Thr Pro Gln Leu Thr Lys Asn Ala Gly Val Leu
330                 335                 340                 345

ACA TGC TCT CTC TCT AAA CGT TGT TGATGATGCA TATATTCTAG CATGTTGTAC     1217
Thr Cys Ser Leu Ser Lys Arg Cys
                350

TATCTAAATT AACATCAACA AGAAAATATA TCATGATGTA TATTGTTGTA TTGATGTCAA     1277

AATAAAAATG TATCTTTTAC TATTAAAAAA AAAAATGATC GATCGGACGG TACCTCTAGA-3'  1337
```

Fig. 4b

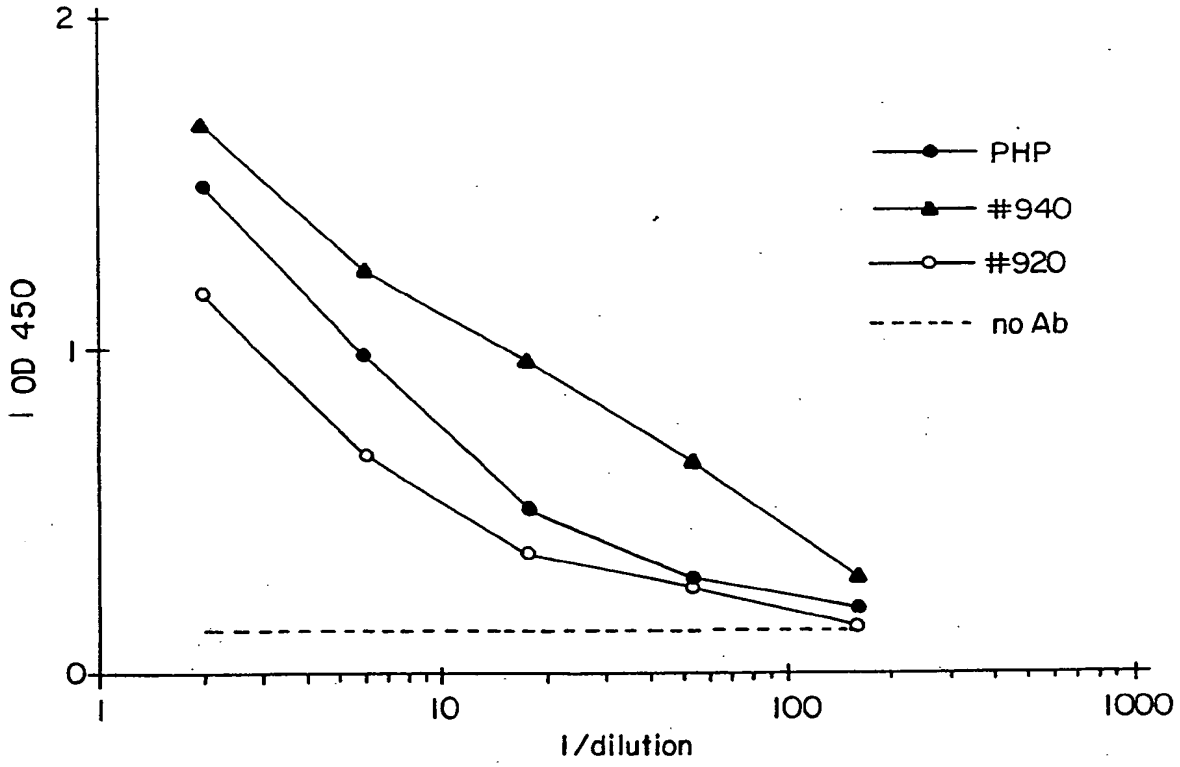

2
1
0
OD 450
1
10
100
1000
1/dilution
PHP
#940
#920
no Ab

Fig. 5a

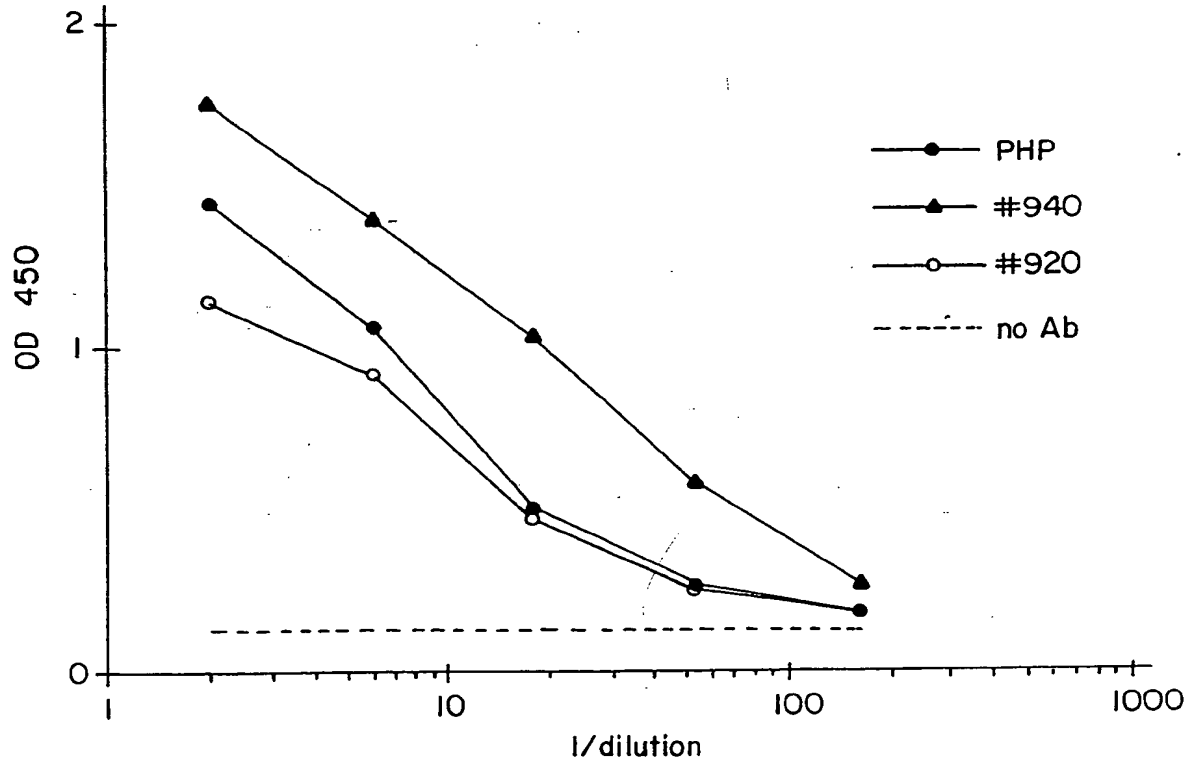

2
1
0
OD 450
1
10
100
1000
1/dilution
PHP
#940
#920
no Ab

Fig. 5b

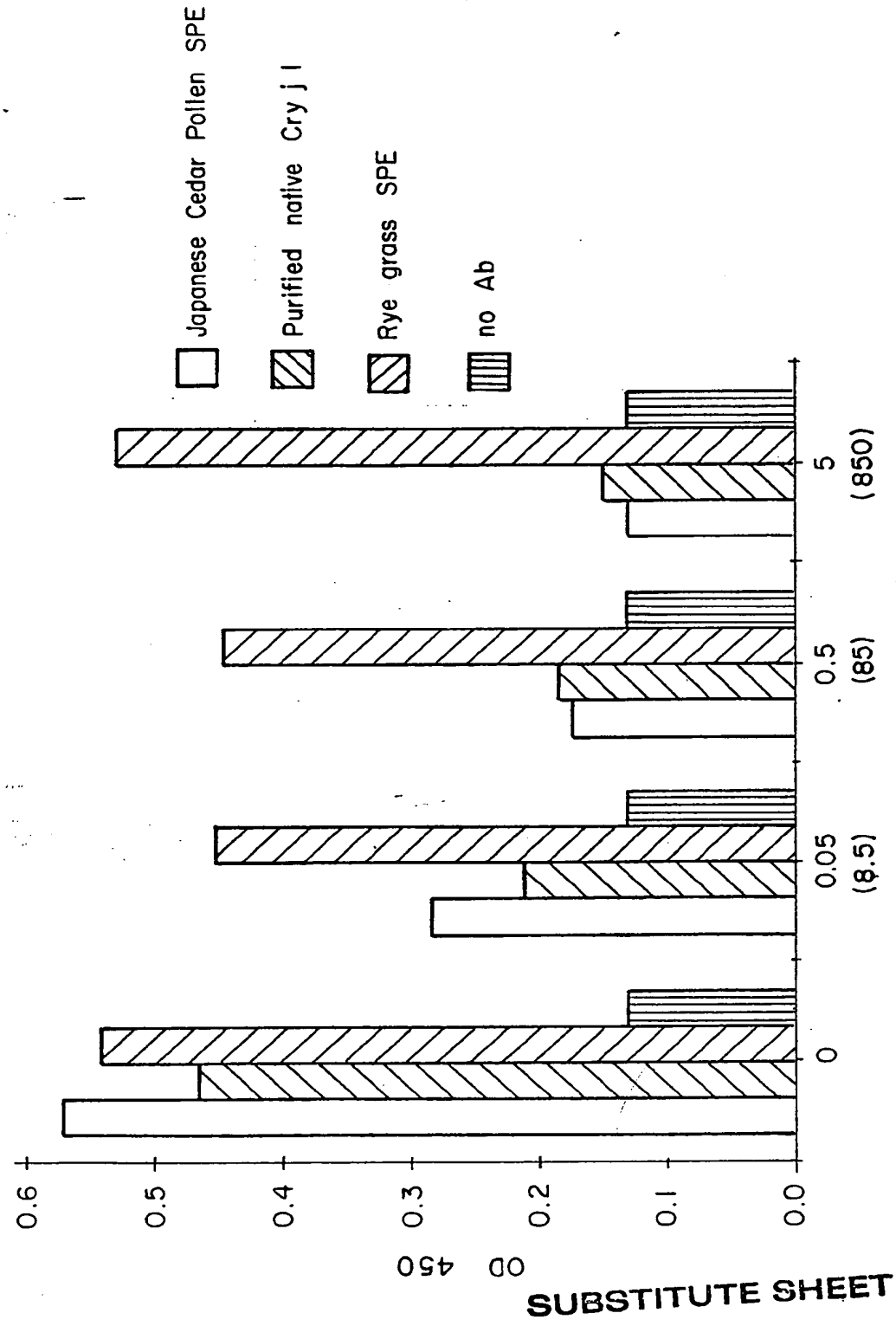

Japanese Cedar Pollen SPE
Purified native Cry j I
Rye grass SPE
no Ab
OD 450
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0
0.05 (8.5)
0.5 (85)
5 (850)
concentration of protein in ug/ml

Fig. 6

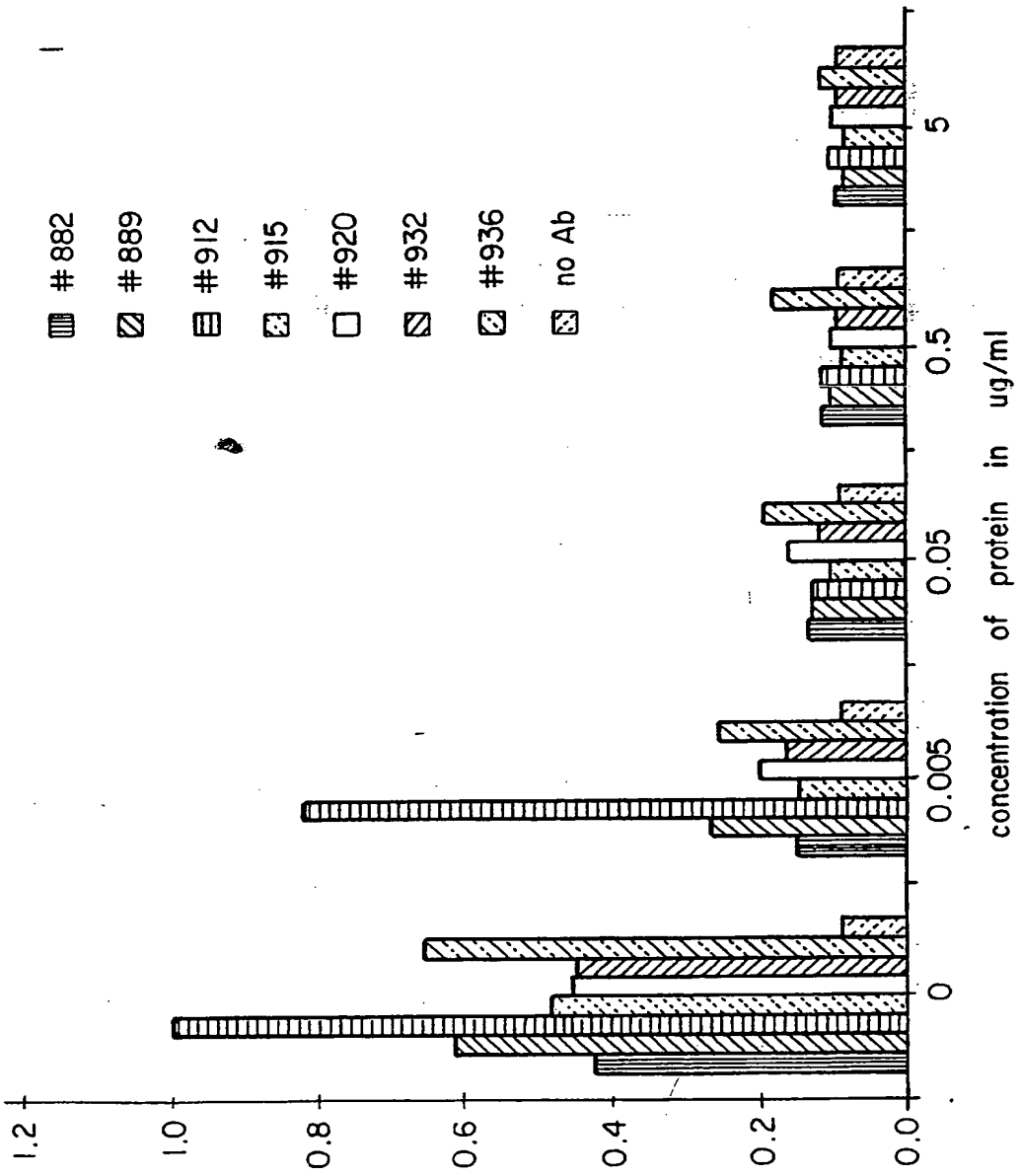
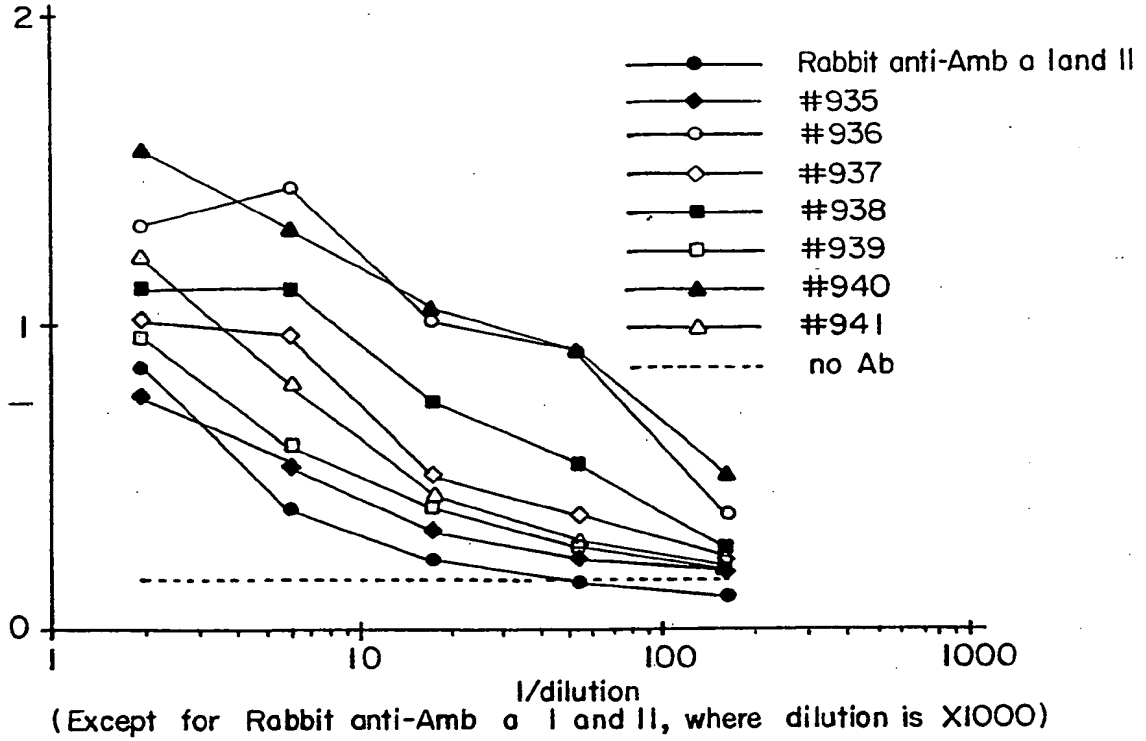
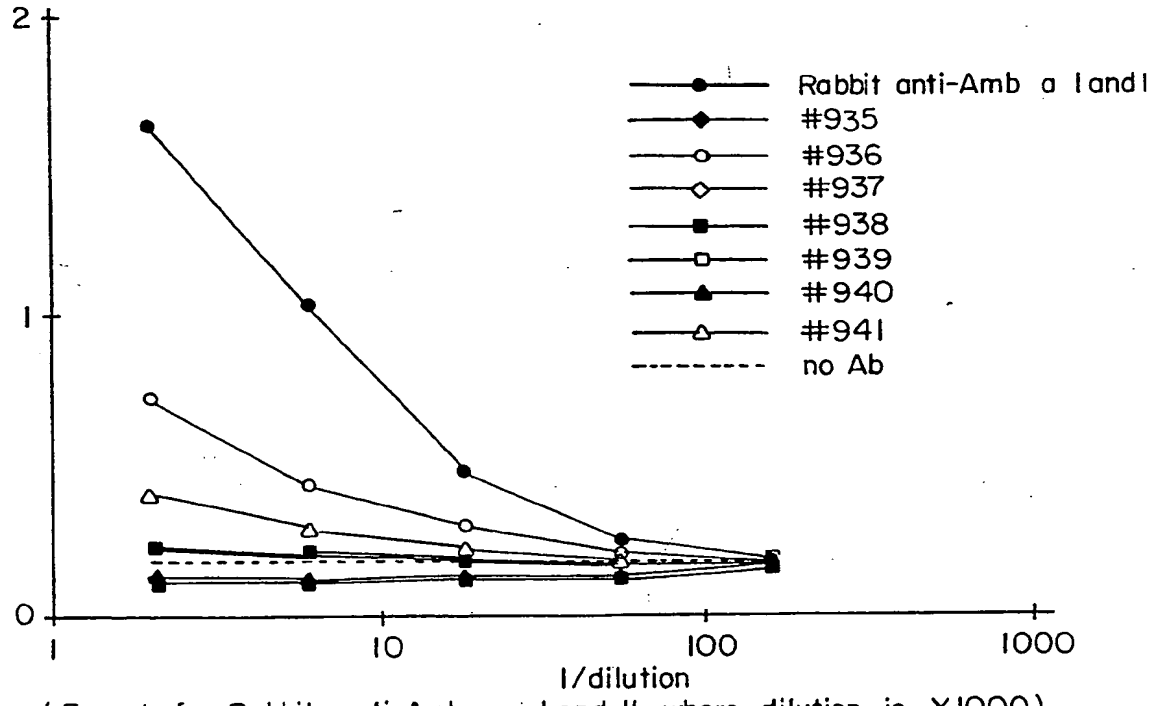

#882
#889
#912
#915
#920
#932
#936
no Ab
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0
0.005
0.05
0.5
5
concentration of protein in ug/ml

Fig. 7



(Except for Rabbit anti-Amb a I and II, where dilution is X1000)

Fig. 8a



(Except for Rabbit anti-Amb a I and II, where dilution is X1000)

Fig. 8b

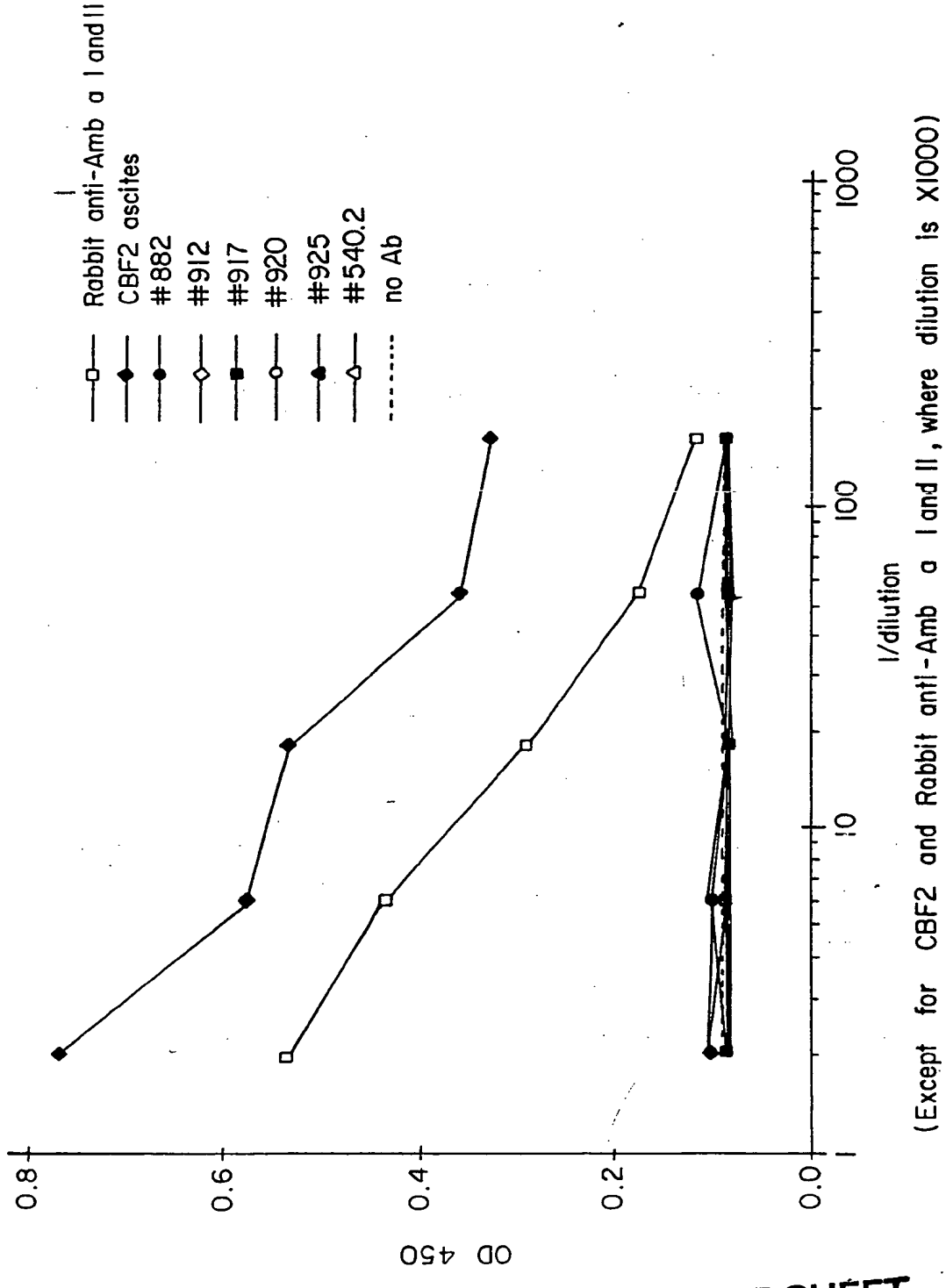

Rabbit anti-Amb a I and II
CBF2 ascites
#882
#912
#917
#920
#925
#540.2
no Ab
OD 450
0.0
0.2
0.4
0.6
0.8
1
10
100
1000
1/dilution
(Except for CBF2 and Rabbit anti-Amb a I and II, where dilution is X1000)

Fig. 9

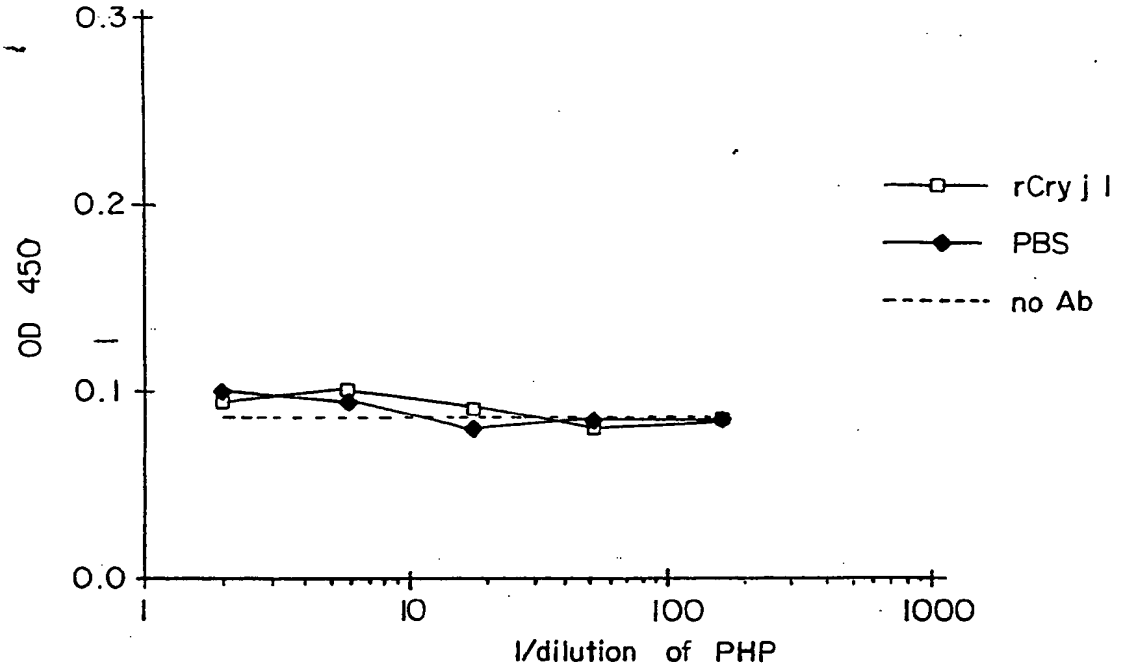


Fig. 10a

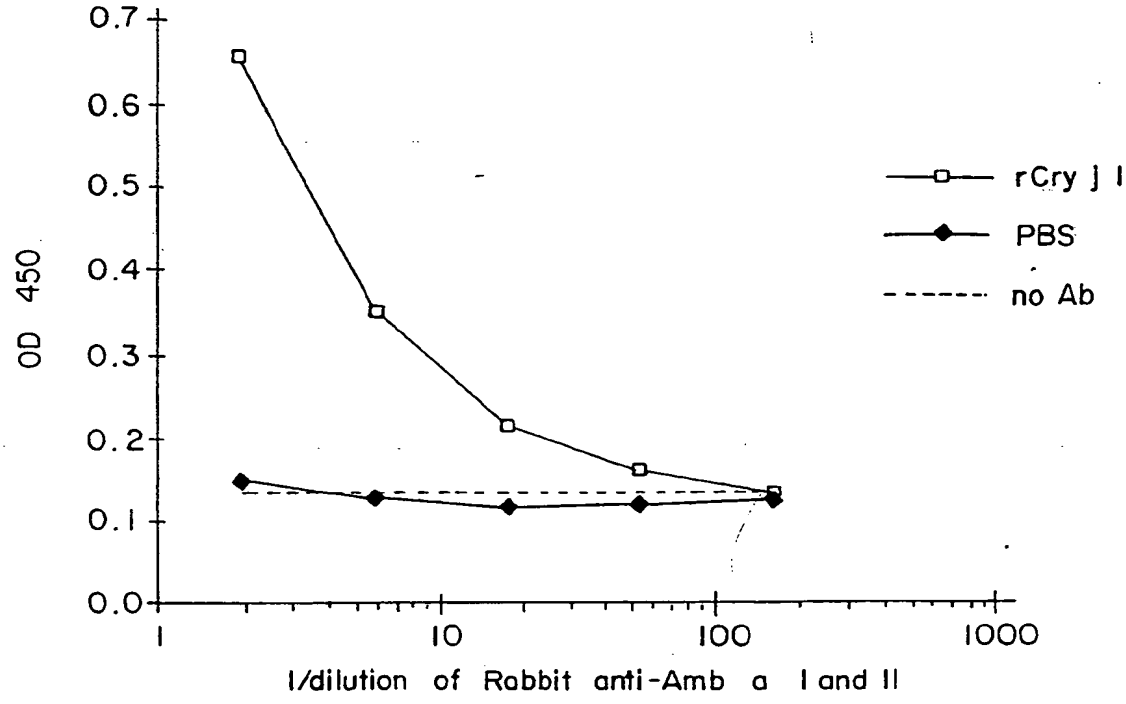


Fig. 10b

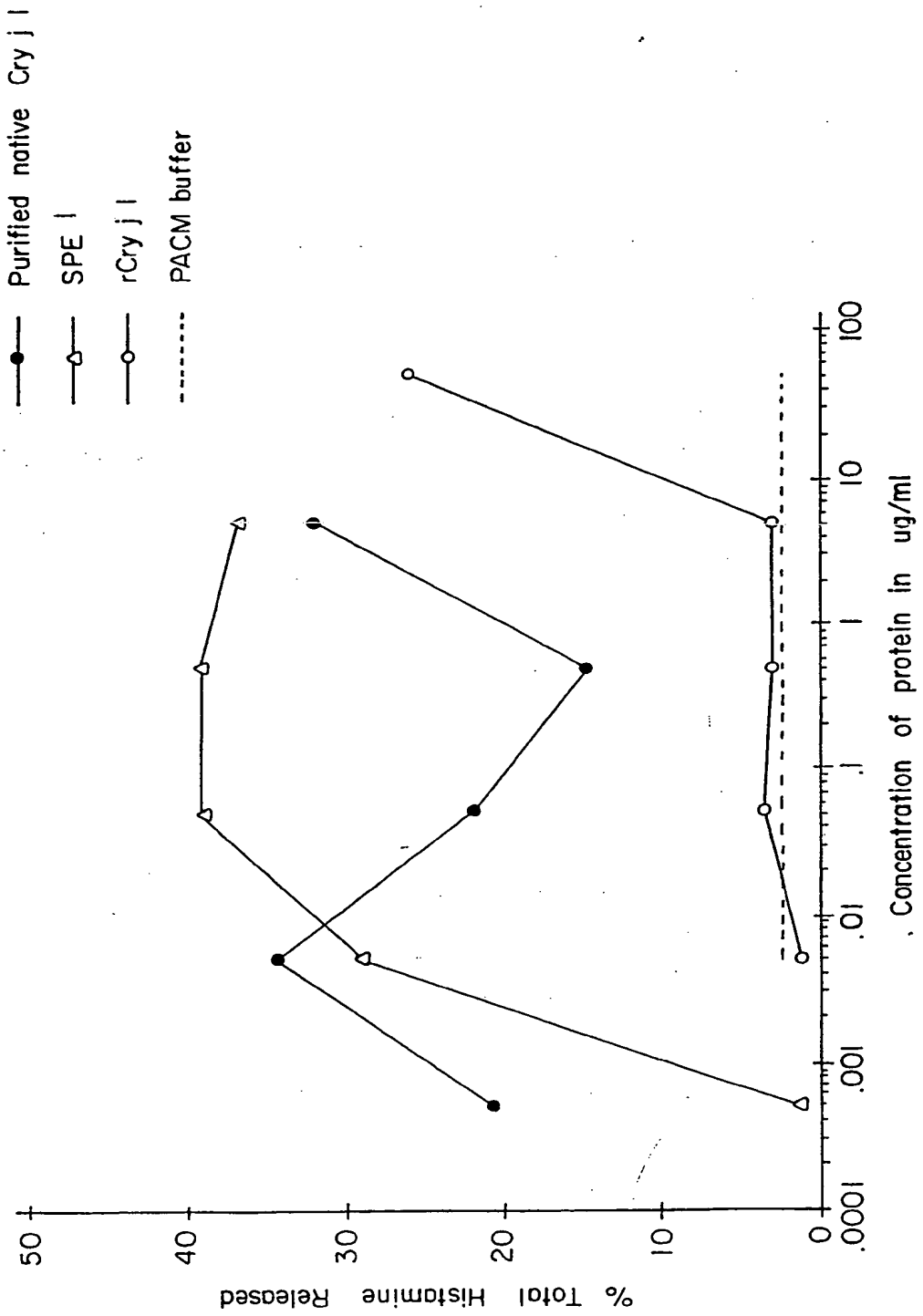

Purified native Cry j I
SPE I
rCry j I
PACM buffer
% Total Histamine Released
0
10
20
30
40
50
.0001
.001
.01
.1
1
10
100
Concentration of protein in ug/ml

Fig. 11

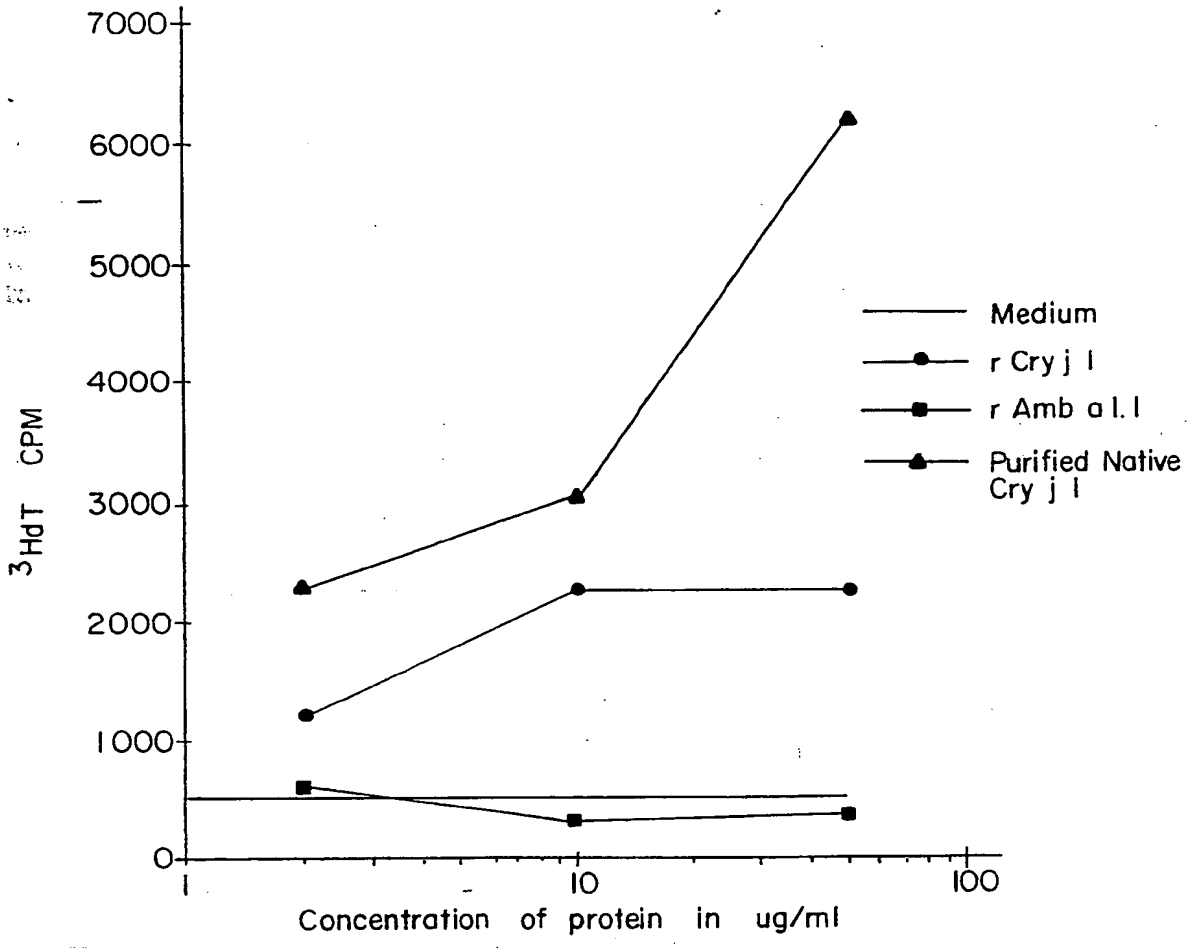

7000
6000
5000
4000
3000
2000
1000
0
^{3}HdT CPM
1
10
100
Concentration of protein in ug/ml
Medium
r Cry j I
r Amb a I. I
Purified Native
Cry j I

Fig. 12

INTERNATIONAL SEARCH REPORT

International Application No PCT/US 92/05661

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all)[6]

According to International Patent Classification (IPC) or to both National Classification and IPC

IPC5: C 07 K 15/10, A 61 K 39/36

II. FIELDS SEARCHED

Minimum Documentation Searched[7]

Classification System	Classification Symbols
IPC5	C 07 K; A 61 K

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in Fields Searched[8]

—

III. DOCUMENTS CONSIDERED TO BE RELEVANT[9]

Category *	Citation of Document,[11] with indication, where appropriate, of the relevant passages[12]	Relevant to Claim No.[13]
X	FEB, vol. 239, No. 2, November 1988, Madoka Taniai et al.: "N-terminal amino acid sequence of a major allergen of Japanese cedar pollen (Cry j I) ", see page 329 - page 332 see especially table 2	5,8,11-36,43
A	EP, A1, 0416816 (KABUSHIKI KAISHA HAYASHIBARA SEIBUTSU KAGAKU KENKYUJO) 13 March 1991, see the whole document	1-37, 43
A	Patent Abstracts of Japan, Vol 12, No 433, C543, abstract of JP 63-159324, publ 1988-07-02 LION CORP ET AL.	1-37, 43

* Special categories of cited documents: [10]

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance, the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance, the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search	Date of Mailing of this International Search Report
1st October 1992	26 OCT 1992
International Searching Authority	**Signature of Authorized Officer**
EUROPEAN PATENT OFFICE	Mikael G:son Bergstrand

III. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)

Category *	Citation of Document, with indication, where appropriate, of the relevant passages	Relevant to Claim No
X	EP, A1, 0308147 (KABUSHIKI KAISHA HAYASHIBARA SEIBUTSU KAGAKU KENKYUJO) 22 March 1989, see especially claim 2	5,8,11-36,43

ernational application No.

INTERNATIONAL SEARCH REPORT

PCT/US92/05661

B x I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.: 38-42
because they relate to subject matter not required to be searched by this Authority, namely:
See PCT Rule 39.1(iv)
Methods for treatment of the human or animal body by surgery or therapy, as well as diagnostic methods

2. [ ] Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

—

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. [ ] As all searchable claims could be searches without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

ANNEX TO THE INTERNATIONAL SEARCH REPORT ON INTERNATIONAL PATENT APPLICATION NO. PCT/US 92/05661

SA 62104

This annex lists the patent family members relating to the patent documents cited in the above-mentioned international search report.
The members are as contained in the European Patent Office EDP file on 28/08/92
The European Patent office is in no way liable for theseparticulars which are merely given for the purpose of information.

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
EP-A1- 0416816	13/03/91	JP-A-	3093730	18/04/91
		US-A-	5073628	17/12/91
EP-A1- 0308147	22/03/89	JP-A-	1156926	20/06/89
		US-A-	4939239	03/07/90